Subject:- **Cyber Security Threat of ChatGPT (Advisory No. 36)**

Context. Advisory No. 31 on the subject with emphasis on Cyber Security challenges/aspects was shared by NTISB on 19th June, 2023. Recently, a breach of around 100,000 Chat GPT user accounts on dark web through an information stealing malware (Raccoon, Vider, Redline) is reported. The report also highlights one of the major challenges of Al driven projects (including ChatGPT); the sophistication of Cyber-attacks. Precautionary measures and cautious use of ChatGPT (at organizational and individual level) are illustrated in ensuing paras.

2. **Trend - Chat GPT User Accounts**. Globally, many organizations are integrating ChatGPT and other Al powered APIs into their operational flows/information systems. ChatGPT accounts signify the importance of Al-powered tools along with the associated Cyber risks as it allows users to store conversations. In case of breach, access of a user account may provide insight into proprietary information, area of interest/ research, internal operational/business strategies, personal communications and software code etc.

3. **Precautionary Measures**

a. **Users**

(1) Do not enter sensitive data into ChatGPT. If essential, ensure to disable the chat saving feature from the platform's settings menu or manually delete those conversations as soon as possible.

(2) Use a malware free/screened system for ChatGPT. An infected system (with information stealer malware) may take snap screenshots or perform keylogging, leading to a data leak.

(3) ChatGPT/other Al-powered tools and APIs must not be used by users handling extremely sensitive data. Masking of critical information/ dummy data may be utilized where absolutely essential.

b. **Organizations**. Through best practices, organizations can ensure that ChatGPT is used securely and the data is protected. It is also important to note that Al technology is constantly evolving. The key to protection may be that organizations must stay up-to-date with the latest security trends. Few best practices (but not limited to) are as follows:

(1) **Conduct Risk Assessment**. Before use of ChatGPT, conduct a comprehensive risk assessment to identify any potential/exploitable

vulnerabilities. This will help organizations to develop a plan to mitigate risks and ensure that their data is protected.

(2) **Use Secure Channels**. To prevent unauthorized access to ChatGPT, use secure channels to communicate with the chatbot. This includes using encrypted communication channels and secure APIs.

(3) **Mechanism to Monitor Access**. It is important to monitor who has access to ChatGPT. A mechanism be ensured that access is granted only to authorize individuals. This can be achieved by implementing strong access controls and monitoring access logs.

(4) **Implement Zero-Trust Security**. Zero-trust security (an approach that assumes that every user and device on a network is a potential threat) be adopted. This means that access to resources should be granted only on a need-to-know basis followed by strong authentication mechanism.

(5) **Train the Employees**. Employees be trained on use of ChatGPT and the potential risks associated with its use. The employees do not share sensitive data with chatbot and are aware of the potential threat of social engineering attacks.