Quais cuidados devem ser observados ao capturar dados de um site?

* Abrir o aquivo robots.txt presente nos sites para saber quais caminhos tem ou não acesso permitido.
* Não fazer scrapers antes de conferir se a página possui direitos autorais.
* Não capturar dados sensíveis, mesmo que sejam públicos.
* Não fazer uma grande captura a ponto de derrubar o servidor do site.