

DATA

ANONYMIZATION

AGENDA

- ONE** obvious need
- TWO** security concerns
- THREE** possible solution
- FOUR** demo and todos

NEED

- * devs need to **work** with **real production data**
- * because we have to **debug** and **test** applications

CONCERNS

- * but we **don't need** to access **sensitive data**
- * because it puts **users personal records** at risk

RISKS

- * extended **attack surface**
- * **lost or stolen** device
- * otherwise **insecure environment**
- * **remote** team member

SOLUTION

- * common approach for this is to **take a production database dump**

- * and **anonymize it**

WHITELISTING

- * **comprehensive** approach
- * takes **longer to run**
- * might make **application harder to test**

BLACKLISTING

- * is selective

- * new or unidentified data might slip through

DEMO

TODO

- * **complete script**
- * **integrate** with our workflow
- * improve to **anonymize data** dumps **on the fly**

QUESTIONS

LINKS

- * sunitSPACE.blogspot.ie/2012/09/data-anonymization.html
- * sunitSPACE.blogspot.ie/2012/09/data-anonymization-blacklist-whitelist.html
- * github.com/sunitparekh/data-anonymization
- * github.com/Barzahlen/tidus