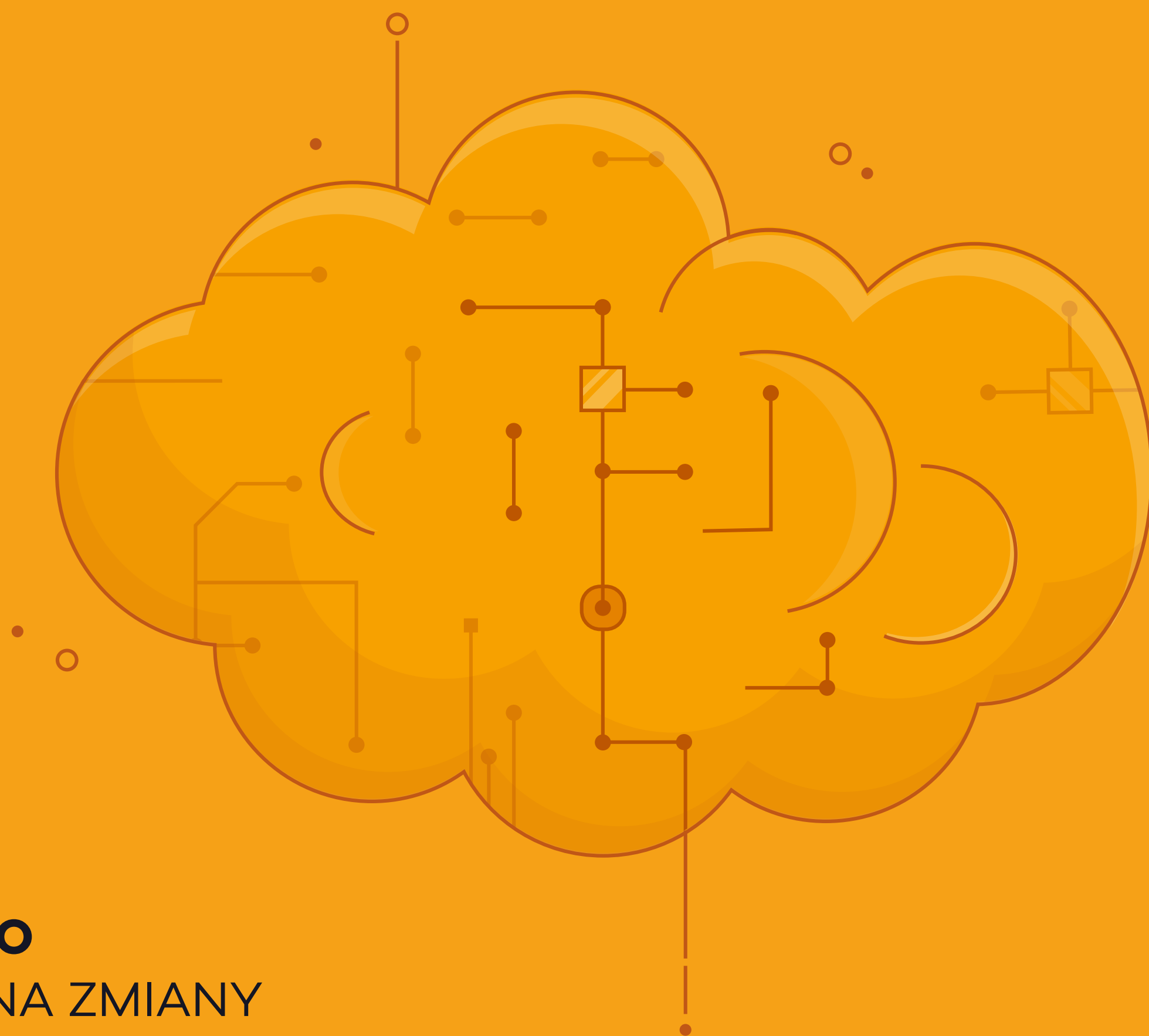


AI ACT WEEKLY #14
DATA GOVERNANCE
W SYSTEMACH AI
WYSOKIEGO
RYZYKA



BĄDŹ NA BIEŻĄCO
PRZYGOTUJ SIĘ NA ZMIANY

Praktyki zarządzania danymi

Systemy wysokiego ryzyka, które wykorzystują techniki obejmujące trenowanie modeli AI z wykorzystaniem danych (treningowych / walidacyjnych / testowych), powinny być trenowane tylko na takich danych, które spełniają określone w AI Act kryteria jakości.



Dane powinny podlegać dobrym praktykom zarządzania danymi. Należą do nich m.in.:

- odpowiednie decyzje projektowe,
- sposoby zbierania danych i ustalania ich pochodzenia,
- przygotowywanie danych – czyszczenie, komentowanie, etykietowanie, aktualizacja, wzbogacanie,
- badanie stronniczości i zapobieganie dyskryminacji,
- określenia istotnych luk lub braków w danych.

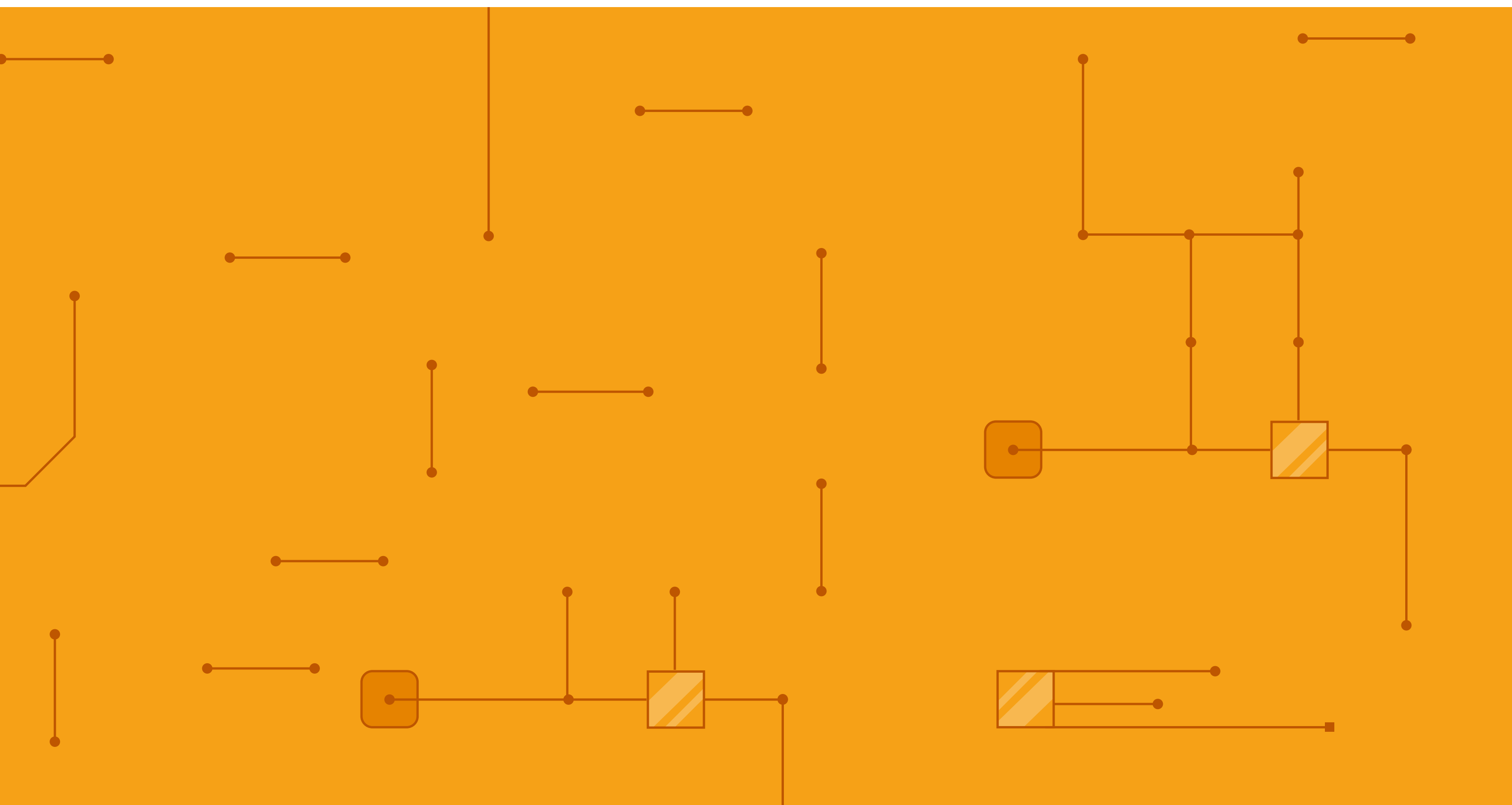


Jakość i reprezentatywność danych

- Zbiory danych powinny być adekwatne, wystarczająco reprezentatywne, w jak największym stopniu wolne od błędów i kompletne z punktu widzenia przeznaczenia.
- Dane powinny się charakteryzować odpowiednimi właściwościami statystycznymi np. w odniesieniu do osób lub grup osób, wobec których ma być stosowany system.
- Zbiory danych powinny uwzględniać kontekst geograficzny, kulturowy, behawioralny i funkcjonalny w którym ma być wykorzystywany system.



W przypadkach rozwoju systemów AI wysokiego ryzyka niewykorzystujących technik obejmujących trenowanie modeli AI, kryteria jakości i dobre praktyki dotyczące data governance stosujemy do danych testowych.



Potrzebujesz wsparcia przy wdrożeniu AI Act w Twojej organizacji? Zapraszamy do kontaktu!



MICHAŁ JACKOWSKI

M.JACKOWSKI@DSK-KANCELARIA.PL



ALEKSANDER ZIELIŃSKI

A.ZIELINSKI@DSK-KANCELARIA.PL
