

Projekt RRP5 VeMoInd

D5.2 Poročilo o primernosti tehnične opreme in možnih načinih integracije jezikovnih in govornih tehnologij v proizvodnih okoljih

Avtor dokumenta: Andrej Planina, Špica International d.o.o.

Datum: 16. avgust 2024

1. Povzetek.....	2
1.1. Uvod	2
1.2. Namen dokumenta.....	2
1.3. Povezava z drugimi deli projekta	2
2. Izbrani pristop.....	2
3. Rezultati	3
4. Zaključek	3

1. Povzetek

1.1. Uvod

Naloga, ki smo si jo zadali v okviru projekta RRP5 VeMoInd, je **Prilagoditev LLM za uporabo v industrijskih okoljih**. Znotraj tega projekta je bil določen cilj **C5.2 Preučiti in preskusiti tehnično opremo in tehnične možnosti za vpeljavo jezikovnih in govornih tehnologij v industrijskih okoljih**.

Ena od nalog projekta je »izbrati opremo (slušalka in mikrofona), ki bo primerna za uporabo glasovne komunikacije med IT sistemi in zaposlenimi v proizvodnih podjetjih«. Cilj je bil zagotoviti tako opremo, ki omogoča učinkovito in zanesljivo komunikacijo v hrupnih industrijskih okoljih.

Druga naloga v okviru projekta je »raziskati možne načine integracije jezikovnih in govornih tehnologij v proizvodnih okoljih«, kjer smo identificirali tiste IT rešitve, s katerimi bi bilo potrebno s pomočjo govornih tehnologij povezati zaposlene in raziskali načine za povezavo z njimi.

1.2. Namen dokumenta

Namen pričujočega dokumenta je povzeti vse ugotovitve, do katerih smo prišli tekom raziskave na temo primernosti tehnične opreme za uporabo jezikovnih in govornih tehnologij v proizvodnih okoljih, kar predstavlja en del predvidenega rezultata D5.2. Dokument ponuja pregled metodologije, rezultatov in podaja priporočila glede izbora opreme za glasovno komunikacijo v proizvodnih okoljih.

Drugi namen dokumenta je prikazati najbolj pogoste IT rešitve, s katerimi bi bilo potrebno povezati glasovne tehnologije v proizvodnih okoljih in prikazati možnosti za tehnične načine povezave.

1.3. Povezava z drugimi deli projekta

Povezava te raziskave z drugimi deli projekta je v tem, da iščemo tako opremo, ki bi tudi zaposlenim v industrijskem okolju omogočala dovolj učinkovito uporabo velikih jezikovnih modelov in to v govoru, ki je zaposlenim blizu oz. jim je to naravni govor. Tako opremo, ki bi jo skupaj z drugimi partnerji lahko vključili v celoviti rešitev za podporo glasovni komunikaciji preko velikih jezikovnih modelov, v industrijskih okoljih z vsemi njihovimi posebnostmi.

Dodatno dokument prikazuje, kako je možno govorno tehnologijo povezati v obstoječe IT sisteme v proizvodnji, kar je pomembno za projektno nalogo »prototipna integracija sistema za govorno komunikacijo v eno od Špica rešitev«

2. Izbrani pristop

Pri aktivnostih smo uporabili naše dosedanje izkušnje z glasovnim vodenjem dela v proizvodnih okoljih. Pri konkretni raziskavi smo za prepoznavanje govora in njegovo pretvorbo v besedilo (STT, Speech to Text) uporabili prepoznavanje omejene količine vnaprej definiranih izrazov (cca. 100

izrazov). Vsak izraz je bil dolg 1-3 besede, vseboval je tudi številke 0 - 100 in nismo uporabili prepoznavanja splošnega teksta, ker ta del rešitve trenutno še ni bil razvit. Pri generiranju zvoka iz teksta (TTS, Text to Speech) smo uporabili daljše izraze, dolge do nekaj besed, kar je tipičen način komunikacije pri trenutnih rešitvah za glasovno vodenje dela v proizvodnih okoljih, predvsem pri interni logistiki in v skladiščih.

Izbrali smo pristop, ki nam je počasi ožil možne rešitve, od splošnih zahtev, preko laboratorijskih preizkusov, do testiranja z uporabniki na terenu. Sledili smo trem fazam:

1. Priprava širokega nabora razne možne opreme glede na tehnične lastnosti opreme
2. Izbira in izločanje opreme, ki ne zadošča minimalnim tehničnim zahtevam projekta.
3. Testiranje nekaj različnih kompletov opreme s konkretnimi uporabniki na terenu in zbiranje njihovega mnenja

Opisanemu pristopu smo strogo sledili in nam je prinesel zadovoljive rezultate, za katere ocenjujemo, da so argumentirani in dovolj dobri za uporabo v drugih delih projekta, pri drugih projektnih partnerjih in pri drugih naših projektih.

Do rezultatov smo prišli večinoma s svojimi lastnimi aktivnostmi in s posvetovanjem in angažmajem drugih projektnih partnerjev, predvsem Fakultete za računalništvo in informatiko ter podjetja Vitasys. Fakulteta je pomagala predvsem glede uporabe in integracije z velikimi jezikovnimi modeli, Vitasys predvsem glede odpravljanja šumov in hrupa.

V okviru raziskave na temo integracije v obstoječe IT sisteme v proizvodnji smo uporabili naše dosedanje izkušnje pri uvajanju IT rešitev v proizvodnjo, predvsem znanje o integraciji v obstoječe sisteme in pričakovanja in posebnost uporabnikov – zaposlenih. Do rezultatov smo prišli s pomočjo pogovorov s Špica svetovalci, arhitekti rešitev in razvijalci, pa tudi s pomočjo raziskave literature, pogovorov z našimi strankami in partnerji. Pridobljene rezultate na to temo bomo uporabili pri izdelavi prototipa uvedbe govorne tehnologije v eno od Špica rešitev.

3. Rezultati

--- POSLOVNA SKRIVNOST ---

4. Zaključek

Za uporabo glasovne tehnologije v proizvodnih okoljih so kot osebna oprema primerne slušalke z mikrofonom proizvajalca Jabra, model BlueParrot B450-XT. Oblika slušalk lahko bistveno vpliva na varnost, počutje in učinkovitost delavcev, zato je izboru opreme potrebno posvetiti precej pozornosti. Upoštevati je potrebno tudi subjektivna mnenja delavcev, saj bodo oni v zelo tesnem stiku z opremo vsak dan, tudi po 8 ur hkrati.

D5.2 Poročilo o primernosti tehnične opreme in možnih načinih integracije

Kot najboljše orodje za integracijo govornih tehnologij v obstoječe IT sisteme je najbolj primerna uvedba posebne middleware rešitve, ki izvaja pretvorbo besedilo<->govor in hkrati še pretvorbo dokument<->opravilo in optimizacijo opravil. Primerna je tudi uporaba take ERP/WMS/MES rešitve, ki ima te pretvorbe in optimizacije vgrajene že v sebi, vendar so take rešitve izredno redke.