Konfiguracja zapowiedzi z wykorzystaniem SSML

SSML (Speech Synthesis Markup Language) to język znaczników służący do kontrolowania sposobu odczytywania zapowiedzi przez syntezator mowy. Dzięki SSML można wpływać na tempo czytania, wysokość głosu, głośność, dodawać pauzy, definiować wymowę wyrazów oraz wskazywać sposób interpretacji dat, godzin i liczb podczas generowania zapowiedzi w Platformie Głosowej Grupa.

Poniżej znajdziesz przykładowe zastosowania znaczników SSML w komunikatach głosowych. Przykłady pokazują najczęstsze sytuacje spotykane w zapowiedziach IVR, infoliniach oraz automatycznych komunikatach dla klientów.

- **Użycie kilku różnych języków** – przygotuj jedno powitanie odtwarzane kolejno w kilku językach, aby ułatwić obsługę klientów zagranicznych.
- **Przerwy w komunikacie** – dodaj krótkie pauzy pomiędzy ważnymi informacjami, aby komunikat był bardziej naturalny i łatwiejszy do zrozumienia.
- **Wyróżnienie treści w komunikacie** – zaakcentuj ważne informacje, które powinien usłuszeć klient.
- **Przyspieszenie odczytu długiego komunikatu** – skróć czas odczytu obowiązkowych informacji o nagrywaniu rozmów lub przetwarzaniu danych osobowych.
- **Spowolnienie odczytu ważnych danych** – odczytaj wolniej numer PESEL, numer recepty lub inne dane, które klient powinien przygotować przed rozmową.
- **Naturalne odczytywanie dat** – przedstaw datę wizyty, dostawy lub realizacji usługi w naturalnej formie zamiast jako ciąg cyfr.

- **Naturalne odczytywanie godzin** – w czytelny sposób przekaż godziny pracy firmy lub termin umówionego kontaktu.
- **Odczytywanie numerów telefonów** – podaj numer kontaktowy w formie ułatwiającej jego zrozumienie i zapisanie.
- **Odczytywanie kodów i identyfikatorów** – odczytaj numer zgłoszenia, identyfikator klienta lub kod odbioru cyfra po cyfrze.
- **Naturalne odczytywanie wartości liczbowych** – odczytaj liczbę punktów, produktów lub zgłoszeń jako pełny liczebnik, dzięki czemu komunikat będzie brzmiał bardziej naturalnie.

Przykładowe i najczęściej wykorzystywane zapowiedzi

Przykład podstawowej struktury

Każda zapowiedź przygotowana w SSML musi rozpoczynać się znacznikiem `<speak>` i kończyć znacznikiem `</speak>`.

```
<speak>
Dzień dobry. Witamy w naszej firmie.
</speak>
```

Powitanie/zapowiedź w wielu językach

Jeśli Twoja firma obsługuje klientów posługujących się różnymi językami, możesz przygotować jedno powitanie zawierające kilka wersji językowych. Dzięki temu osoba

dzwoniąca od razu usłyszy komunikat w zrozumiałym dla siebie języku.

W poniższym przykładzie ten sam komunikat jest odtwarzany kolejno w języku polskim, angielskim, niemieckim i ukraińskim.

```
<speak>
  <voice name="pl-PL-Standard-A">Witamy w firmie XYZ, dziękujemy za kontakt.
</voice>
  <voice name="en-GB-Standard-A">Welcome to XYZ company, thank you for
contacting us.</voice>
  <voice name="de-DE-Standard-A">Willkommen bei der Firma XYZ, vielen Dank fü
Ihren Anruf.</voice>
  <voice name="uk-UA-Standard-A">Ласкаво просимо до компанії XYZ, дякуємо
звернення.</voice>
</speak>
```

W Polsce najczęściej stosowanymi językami w zapowiedziach wielojęzycznych są: polski (pl-PL), angielski (en-GB), ukraiński (uk-UA), niemiecki (de-DE), rosyjski (ru-RU), francuski (fr-FR), hiszpański (es-ES), włoski (it-IT), czeski (cs-CZ) oraz słowacki (sk-SK).

W atrybucie name znacznika <voice> należy podać identyfikator wybranego głosu, np. pl-PL-Standard-A, en-GB-Standard-A lub uk-UA-Standard-A. Pierwsza część nazwy określa język i region, natomiast ostatnia część, np. Standard-A lub Standard-B, wskazuje konkretny wariant głosu dla wybranego języka.

W większości języków dostępny jest co najmniej jeden głos oznaczony jako Standard-A, jednak liczba dostępnych wariantów głosowych różni się w zależności od języka. Niektóre języki oferują dodatkowe warianty, takie jak Standard-B, Standard-C lub Standard-D, podczas gdy inne udostępniają tylko jeden głos. Przed wykorzystaniem głosu w komunikacie

warto sprawdzić listę dostępnych wariantów dla wybranego języka.

Wstawienie przerwy w komunikacie

Jeśli komunikat jest dłuższy lub zawiera ważne informacje, możesz użyć znacznika <break>, aby dodać krótką pauzę między jego częściami. Dzięki temu zapowiedź brzmi bardziej naturalnie i jest łatwiejsza do zrozumienia przez osobę dzwoniącą.

W poniższym przykładzie znacznik <break> wstawia przerwy trwające odpowiednio pół sekundy i 1 sekundę, co pozwala oddzielić od siebie poszczególne informacje w komunikacie.

```
<speak>
  Witamy w firmie XYZ.
  <break time="500ms"/>
  Dziękujemy za kontakt.
  <break time="1s"/>
  Za chwilę połączymy Cię z konsultantem.
</speak>
```

Długość pauzy można podać w milisekundach (ms) lub sekundach (s). Na przykład 500ms oznacza pauzę trwającą pół sekundy, a 1s pauzę trwającą jedną sekundę.

Wyróżnienie ważnych informacji w komunikacie

Jeśli chcesz zwrócić uwagę osoby dzwoniącej na konkretną informację, taką jak numer opcji w menu, godziny pracy lub ważny komunikat, możesz użyć znacznika <emphasis>. Dzięki

temu wybrane słowo lub fragment zostanie wypowiedziany z większym lub mniejszym naciskiem niż pozostała część komunikatu.

```
<speak>
 Informacja o promocji jest
 <emphasis level="reduced">opcjonalna.</emphasis>
 Termin realizacji zamówienia wynosi
 <emphasis level="moderate">trzy dni robocze.</emphasis>
 Aby połączyć się z konsultantem, wybierz
 <emphasis level="strong">zero.</emphasis>
</speak>
```

W powyższym przykładzie słowo „opcjonalna” zostało wypowiedziane mniej wyraźnie (reduced), informacja o czasie realizacji została delikatnie wyróżniona (moderate), a numer opcji w menu został mocno zaakcentowany (strong).

Przyspieszenie odczytu długiego komunikatu

Jeśli część komunikatu zawiera obowiązkowe informacje, takie jak informacja o nagrywaniu rozmów lub przetwarzaniu danych osobowych, możesz użyć znacznika <prosody>, aby odczytać ją szybciej. Dzięki temu osoba dzwoniąca krócej czeka na połączenie z konsultantem, a wymagane informacje nadal zostają przekazane.

W poniższym przykładzie obowiązkowy komunikat jest odczytywany szybciej dzięki parametrowi rate="fast". Atrybut rate określa tempo odczytu tekstu. Wartość fast powoduje szybsze wypowiadanie komunikatu niż standardowe tempo.

```
```

```
<speak>
  Witamy w przychodni XYZ.
  <prosody rate="fast">
    Informujemy, że rozmowy mogą być nagrywane, a dane osobowe są przetwarzane zgodnie z obowiązującymi przepisami oraz polityką prywatności placówki.
  </prosody>
  Aby umówić wizytę, pozostań na linii.
</speak>
```

Wolniejsze odczytanie ważnych informacji

Jeśli osoba dzwoniąca musi przygotować konkretne dane lub wykonać określoną czynność, możesz użyć znacznika <prosody>, aby odczytać tę część komunikatu wolniej. Dzięki temu informacje będą łatwiejsze do zrozumienia i zapamiętania.

W poniższym przykładzie dane wymagane podczas rozmowy są odczytywane wolniej dzięki parametrowi rate="slow". Atrybut rate określa tempo odczytu tekstu. Wartość slow powoduje wolniejsze wypowiadanie komunikatu, co ułatwia zrozumienie i zapisanie ważnych informacji.

```
<speak>
  Aby przyspieszyć obsługę, przygotuj:
  <prosody rate="slow">
    numer PESEL oraz numer recepty.
  </prosody>
  Za chwilę połączymy Cię z konsultantem.
</speak>
```

Poprawne odczytywanie dat

Jeśli w komunikacie występuje data zapisana w formacie liczbowym, warto użyć znacznika <say-as>, aby została odczytana w naturalny sposób. Dzięki temu osoba dzwoniąca usłyszy pełną datę zamiast ciągu cyfr.

W poniższym przykładzie data zapisana jako 2025-02-26 zostanie odczytana jako „dwudziestego szóstego lutego dwa tysiące dwudziestego piątego roku”.

```
<speak>
 Termin realizacji zamówienia to
 <say-as interpret-as="date">2025-02-26</say-as>
</speak>
```

Atrybut interpret-as="date" powoduje interpretowanie wskazanego tekstu jako daty.

Poprawne odczytywanie godzin

Jeśli chcesz przekazać godziny pracy, termin wizyty lub czas realizacji usługi, możesz użyć znacznika <say-as>, aby syntezator poprawnie odczytał godzinę.

W poniższym przykładzie godzina zapisana jako 16:30 zostanie odczytana jako „szesnasta trzydzieści”.

```
<speak>
 Infolinia jest czynna do godziny
 <say-as interpret-as="time">16:30</say-as>
```

```
</speak>
```

Atrybut interpret-as="time" powoduje interpretowanie wskazanego tekstu jako godziny.

Wyraźne odczytywanie numeru telefonu

Jeśli podajesz klientowi numer kontaktowy, możesz użyć znacznika <say-as>, aby każda cyfra została odczytana w sposób odpowiedni dla numeru telefonu. Dzięki temu numer jest łatwiejszy do zrozumienia i zapisania.

W poniższym przykładzie numer telefonu 123456789 zostanie odczytany jako „jeden dwa trzy cztery pięć sześć siedem osiem dziewięć”.

```
<speak>
 W sprawie reklamacji zadzwoń pod numer
 <say-as interpret-as="telephone">123456789</say-as>.
</speak>
```

Atrybut interpret-as="telephone" powoduje interpretowanie wskazanego tekstu jako numeru telefonu.

Odczytywanie kodów i identyfikatorów cyfra po cyfrze

Jeśli klient musi przepisać kod odbioru przesyłki, numer klienta lub identyfikator zgłoszenia, możesz użyć znacznika <say-as>, aby każda cyfra została odczytana osobno. Dzięki temu łatwiej jest zapisać lub zweryfikować podaną wartość.

W poniższym przykładzie identyfikator 123456 zostanie odczytany jako „jeden dwa trzy cztery pięć sześć”.

```
<speak>
 Twój numer zgłoszenia to
 <say-as interpret-as="digits">123456</say-as>.
</speak>
```

Atrybut interpret-as="digits" powoduje odczytywanie każdej cyfry osobno.

Naturalne odczytywanie wartości liczbowych

Jeśli w komunikacie występują liczby określające liczbę punktów, produktów lub zgłoszeń, możesz użyć znacznika <say-as>, aby zostały odczytane jako pełne liczebniki. Dzięki temu komunikat brzmi bardziej naturalnie.

W poniższym przykładzie liczba zostanie odczytana jako „sto dwadzieścia trzy”, a nie jako pojedyncze cyfry.

```
<speak>
 Na Twoim koncie znajduje się
 <say-as interpret-as="cardinal">123</say-as>
 punktów.
</speak>
```

Atrybut interpret-as="cardinal" powoduje interpretowanie liczby jako liczebnika głównego.

Revision #9

Created 13 July 2026 16:56:20 by Damian Szymanski

Updated 14 September 2026 09:43:52 by Damian Szymanski