



اداره کل تحقیقات و جهاد خودکفایی

درخواست ارائه پیشنهاد

برای پروژه

پژوهش، طراحی و پیاده‌سازی سامانه نرم‌افزاری دادمکاو و متن‌کاو
بر روی داده‌های مرکز پژوهش و سنجش افکار صداوسیما
باقابلیت پیش‌بینی رفتار مخاطبان

نسخه 9

شناسه سند: WEVR 950406

تاریخ آخرین تغییرات: 30/10/96

سازمان صداوسیما جمهوری اسلامی ایران
معاونت توسعه و فن‌آوری رسانه
اداره کل تحقیقات و جهاد خودکفایی

الحمد لله الرحمن الرحیم

1- مقدمه

پیشرفت هر جامعه درگرو رشد و تعالی افراد پویای جامعه است و در این بین نمی توان نقش پررنگ رسانه های صداوسیما را در بهبود و ارتقا این رشد و تعالی نادیده گرفت. امروزه رسانه صداوسیما با وجود اینترنت و فراگیر شدن آن عملاً از حالت یکسویه خارج شده و ماهیت متفاوتی پیدا کرده است. در حال حاضر مخاطبان علاوه بر اینکه مصرف کننده اطلاعات هستند، منبع بالقوه ایجاد منابع اطلاعاتی مفید و مثر ثمر نیز هستند و به همین دلیل می توان از این ارتباط دوسویه استفاده ی بهینه را در جهت کسب دانش مورد نیاز برای بهبود کیفیت تولیدات صداوسیما و حتی سمت دهی این تولیدات داشت. امکان دسترسی به

مجموعه داده های گردآوری شده از طریق پیمایش های مختلف و کاوش در این داده ها و تحلیل و پیش بینی رفتار مخاطبان در آینده، یکی از ملزوماتی است که اخیراً نیاز آن به طور ویژه مدنظر قرار گرفته است.

2- بیان مسئله

افکار سنجی اجرای اقدامات و تلاش های سازمان یافته برای نشان دادن عقاید جامعه نسبت به یک موضوع در یک محل خاص و در یک مقطع زمانی معین می باشد و نتیجه آن نشان می دهد که افراد به چه دلیلی عقیده خاصی را پذیرفته اند و با چه شدت و قوتی از آن پشتیبانی می کنند. بنابراین تحلیل افکار و رفتار مخاطبان سازمان بر مبنای این افکار سنجی های مبتنی بر نظرسنجی ها فعالیتی است که علاوه بر فعالان سیاسی، به منظور جلب آراء مردم و یا شناخت و هدایت افکار عمومی برای برنامه ریزان در جهت تنظیم سیاست ها و خط مشی ها و اندازه گیری میزان رضایت مخاطبان سازمان صداوسیما کاربرد دارد. به عنوان مثال هرچه میزان مطابقت محتوای برنامه های تلویزیونی با ویژگی های مخاطبان موردنظر، نیازها و خواسته های آنان بالاتر باشد، میزان اثربخشی آن محتوا و رضایت مخاطبان از آن برنامه ها نیز بیشتر خواهد بود. یکی از راه های ارزیابی میزان رضایت مخاطبان از محتوای تولید شده در رسانه ملی نظرسنجی هایی است که در این زمینه صورت گرفته است. با توجه به جایگاه مهم رسانه ملی، این درخواست ارائه پیشنهاد¹ به منظور ارائه راه حل برای موارد ذیل تدوین شده است.

- ارائه امکانی برای سنجش میزان اثربخشی محتوای تولید شده توسط سازمان بر روی مخاطبان
- استفاده از نتایج نظرسنجی ها در شناسایی رفتار مخاطبان و آشنایی با نیازها و خواسته های آنها
- استفاده از نتایج نظرسنجی ها در پیش بینی رفتار مخاطبان مثلاً در مناسبت های سیاسی مانند انتخابات
- استفاده از روش ها یا تکنولوژی های جدید و کاربردی در دنیا برای مطالعه و تحلیل رفتار مخاطبان رسانه

صداوسیما

- استفاده از امکان جمع و تحلیل یکپارچه نظرسنجی های صورت گرفته از حوزه های مختلف
 - پیش بینی تعداد مخاطب برای برنامه های خاص
 - پیش بینی زمانی که افراد به تماشای برنامه های خاص میپردازند
 - مخاطبان برنامه های خاص
 - مدل ها و الگوهای حاکم بر رفتار مخاطبان
- همچنین چون این پروژه برای بار اول اجرا می شود در ابتدا نیازمند فعالیت های نیازسنجی، تحلیل و پژوهش می باشد.

¹ RFP Request For Proposal

3- هدف پروژه

در حال حاضر مرکز پژوهش و سنجش افکار گروه های مختلفی دارد که هر گروه به طور جداگانه پرسشنامه های خود را در قالب نظرسنجی های مرتبط با گروهشان انجام می دهند و بعد از تجمیع نتایج پرسشنامه ها، توسط کارشناسان آمار و تحلیلگران، نتایج توسط نرم افزار SPSS تحلیل شده و هر گروه به طور مستقل گزارشات مورد نظر خود را می گیرند. هدفی که در این طرح به دنبال آن هستیم خرید سخت افزار و سامانه برای تحلیل خودکار نظرات افرادی است که در نظرسنجی های صداسیما شرکت نموده اند و به سوالات چندگزینه ای و همین طور سؤال پیشنهاد و نظر آزاد پاسخ داده اند. در این تحلیل ارتباط بین پاسخ های افراد به مختلف مورد بررسی قرار خواهد گرفت. این تحلیل می تواند به صورت کلی برای تمام پرسش شوندگان انجام گیرد و یا بر روی گروه های مشخصی که افراد به آن ها تعلق دارند، انجام شود. به عنوان مثال افراد متعلق به یک منطقه جغرافیایی خاص یا افراد دارای مدرک تحصیلی مشخص. (همه این اهداف در ابتدا در سه حوزه انتخابات، مصرف رسانه ملی و ماهواره مدنظر است):

در این سامانه چند بخش کلی در نظر گرفته می شود که در پیوست فنی شماره یک اطلاعات آن موجود است.

4- شرح وضع موجود و نیازمندی ها و ترسیم وضع مطلوب

عنوان	وضع موجود	وضع مطلوب
گزارشات	فقط گزارشات آماری به طور دستی با استفاده از نرم افزار spss	گزارش های آماری و تحلیلی با استفاده از سامانه
روش تحلیل	توسط تحلیلگر	توسط سامانه و تحلیلگر
پیش بینی رفتار مخاطبان	فقط توسط تحلیلگر با توجه به گزارشات آماری	توسط سامانه و بر مبنای تحلیل های خودکار
استفاده از الگوریتم ها و تکنیک های فن آوری روز برای تحلیل	صرفاً مبتنی بر مهارت تحلیلگر	استفاده از تکنیک ها و الگوریتم های مطرح در داده کاوی مانند دسته بندی، خوشه بندی، هوش مصنوعی و شبکه های عصبی و ...

جدول 1- شرح وضع موجود و وضع مطلوب

5- نیازهای غیر عملکردی²

5-1 امنیت نرم افزار

لحاظ نمودن مؤلفه های سه گانه امنیت اطلاعات شامل: محرمانگی، یکپارچگی و در دسترس بودن

5-2 کار آیی

در تولید و توسعه سامانه باید به کار آیی آن توجه ویژه داشت. ارتباط همزمان کاربران تا 20 ارتباط و با 10GB حجم داده برای تحلیل و سرعت اجرای بهینه در نظر گرفته شود.

5-3 واسط کاربری

پیاده سازی سامانه باید مطابق استانداردها و الگوهای معتبر در حوزه ی GUI و user friendly و با تأیید کارفرما باشد و سامانه قابلیت توسعه پذیری داشته باشند.

6- قلمرو اجرای پروژه

مشتری پروژه: مشتری این پروژه مرکز پژوهش و سنجش افکار صداوسیما می باشد.
قلمرو منطقی و فیزیکی: اطلاعات موجود و فایل های تولید شده داده ای از طریق نظرسنجی های انجام شده به صورت مراجعه حضوری، تلفنی و اینترنتی موجود مرکز داده مرکز پژوهش و سنجش افکار صداوسیما
قلمرو فن آوری: هوش مصنوعی، خوشه بندی، داده کاوی، معماری سخت افزار و شبکه و کلیه ابزارهای موجود جهت آنالیز و تحلیل افکار عمومی

7- ارائه مراحل و فازبندی اجرای پروژه

پیشنهاد دهنده بایست مراحل و فازبندی انجام کار را با جزئیات کامل ارائه کند. توجه به موارد زیر در پیشنهاد ارسالی ضروری است:

- پیشنهاد ارسالی محدود به موارد مشخص شده در این مستند نیست و در صورت داشتن ایده کارآمد و یا ویژگی کاربردی در صورت ارائه ارزیابی و لحاظ خواهد شد.

² Non-functional

- در پیشنهاد ارسالی می بایست منابع موجود مورد استفاده (به همراه لایسنس آن ها) و نرم افزارهای تولیدی شرکت به تفکیک مشخص شود. فراهم کردن محیط اجرا از حیث سیستم عامل و بانک اطلاعاتی به عهده کارفرماست ولی سایر ابزارهایی که پیمانکار در سامانه استفاده می کند باید قابلیت پشتیبانی توسط پیمانکار داشته باشد و چنانچه ابزار دچار مشکل شد جابجایی ابزار به عهده پیمانکار می باشد.

8- شیوه ارزیابی

8-1 آزمون تحمل خرابی

- سامانه باید در مقابل خرابی های عمدی یا غیر عمدی توانایی بازگشت به حالت پایدار را داشته باشد. توصیه می شود رفتار سامانه در مواردی نظیر قطع ناگهانی برق، قطع اتصالات شبکه، اشکالات سیستم عامل، اتمام باتری، عدم تنظیم مناسب پارامترهای داخلی، دستکاری عمدی در سیستم فایل های فیزیکی و... بررسی شود.

8-2 آزمون واسط کاربر

- آزمون واسط کاربر توسط نیروهای سازمان انجام می گیرد و باید به تائید کارفرما برسد.

8-3 آزمون Backup & Recovery سامانه

- سامانه باید بتواند از پایگاه داده خود به صورت دستی و یا زمان بندی فایل پشتیبان ایجاد کرده و در صورت لزوم از این فایل ها برای بازیابی به شرایط پایدار و صحیح استفاده کند.

8-4 آزمون امنیت

- سامانه باید گواهی امنیتی مورد تائید کارفرما را داشته باشد.

8-5 تست سامانه

- هزینه های انجام تست سامانه و مستندات آن به عهده پیمانکار است و پیمانکار باید راهکار خود را دربار فرآیند تست و تحویل گیری جهت تائید سازمان اعلام نماید.

9- زمان و شرایط پیاده سازی طرح

- در پایان پروژه، کلیه اسناد مربوط به نصب و راه اندازی و راهنمای سامانه در اختیار سازمان صداسیما قرار می گیرد.

- 20 نفر از کارشناسان و بهره‌برداران از سامانه موضوع قرارداد در سازمان صداسیما در کلیه بخشهای فنی و مدیریتی سامانه تحت آموزش شرکت ارائه‌دهنده سامانه قرار میگیرند.
- تعهدات لازم از شرکت برای پشتیبانی از سامانه در طول یک دوره بهره‌برداری پنج‌ساله اخذ خواهد شد و شرکت طراح موظف به ارائه اسناد (SLA) برای پشتیبانی پنج‌ساله از سامانه موضوع قرارداد (در قبال اخذ هزینه) میباشد. سازمان صداسیما برای نحوه بهره‌برداری از این سامانه مختار خواهد بود.
- سازمان صداسیما در عقد قرارداد با هر شخص یا شرکتی به‌غیر از شرکت مرجع برای توسعه و به‌روزرسانی نرمافزار مختار است و پس از واگذاری شرکت حق هیچگونه ادعای مالکیت مادی یا معنوی نسبت به سامانه در هیچیک از اجزای اصلی و زیرسامانه‌ها نخواهد داشت.
- اخذ تعهدات پشتیبانی از شرکتهای ثالث که تأمین‌کننده بخشهای مدولار نرمافزاری و سختافزاری سامانه هستند (داخلی/خارجی) و نقل و انتقال آن به سازمان صداسیما و یا نماینده آن سازمان به عهده شرکت می‌باشد.
- در صورت عدم موفقیت شرکت در پیاده‌سازی در هر مرحله، سازمان صداسیما نسبت به قطع همکاری و دریافت ضرر و زیان و خسارات وارده خود از محل ضمانتهای شرکت اقدام خواهد نمود.

10- ملاحظات خاص

- 1 کلیه موارد ذکر شده در بند شرح خدمات سامانه و زیرمجموعه‌ها در مدت اجرای طرح مورد ارزیابی دقیق کارشناسان سازمان صداسیما قرار خواهد گرفت.
- 2 پس از تحویل کامل پروژه، پیمانکار ملزم به ارائه سمیناری در سازمان خواهد بود.

11- برنامه زمانی مورد انتظار پروژه

- برنامه زمانی پروژه از زمان عقد قرارداد به مدت 11 ماه در نظر گرفته شده است که می‌بایست در چهار فاز به اتمام برسد.
- فاز شروع: ارائه گزارشات نیازسنجی که در پیوست فنی 2 جزئیات آن آورده شده است.
- فاز 1: تهیه زیرسیستم‌هایی از سامانه که مربوط به ورود داده‌های پیمایش شده در نظرسنجی‌ها و تجمیع آن‌ها در پایگاه داده‌ی یکپارچه و برقراری ارتباط به نظرسنجی‌های مرتبط و تهیه داشبورد ارائه گزارشات آماری مختلف توضیح داده شده در پیوست فنی
- فاز 2: تهیه زیرسیستم‌های پیش‌پردازش داده‌ها و بررسی همبستگی میان داده‌ها، ارائه الگوهای تکرار شونده پنهان در نتایج نظرسنجی‌ها، پیش‌بینی رفتار افراد بر اساس الگوهای استخراج شده، یافتن افراد مشابه هم در جامعه و بیان مهم‌ترین ویژگی‌های آن‌ها، دسته‌بندی داده‌ها بر اساس ویژگی‌های مشابه، پیش‌پردازش متون نظرات آزاد، استخراج موضوعات اصلی

متون به همراه تحلیل آماری آنها، استخراج موضوعات مرتبط با گروه های مختلف شرکت کننده، تحلیل افکار مثبت و منفی حول هر موضوع و ...

فاز 3: تهیه نسخه نهایی و جامع از سامانه

فاز 4: نصب و تست و آزمایش

- در انتهای هر فاز ارائه یک گزارش مفصل و یک گزارش خلاصه مدیریتی الزامی است.

12- چگونگی آموزش

آموزش های راهبری و کاربری برای کارشناسان ذیربط در محل کارفرما باید به صورت کارگاهی یا سمیناری ارائه گردد. تأمین فضا و امکانات آموزش بر عهده کارفرما می باشد. آموزش به صورت 5 جلسه 4 ساعته برای مدیران سامانه و 3 جلسه 3 ساعته برای کاربران عمومی سامانه مورد نیاز می باشد.

13- نصب و پشتیبانی سامانه

پیمانکار موظف است گزارش هر بخش از سامانه را جداگانه ارائه و پس از تست، آموزش و استقرار سامانه تحویل موقت انجام و پس از راه اندازی مورد توافق یک ماهه، سامانه به طور کامل تحویل نهایی گردد.

همچنین مستندات پروژه باید به طور کامل ارائه گردد. کلیه مستندات باید به صورت مکتوب در قالب دفترچه های سیمی یا شیرازه شده ارائه گردد و فایل الکترونیکی قابل ویرایش آنها نیز تحویل دستگاه اجرایی گردد.

در ضمن چنانچه بعد از تست سامانه، باگ هایی (چه امنیتی چه عملکردی) در سامانه موجود بود، پیمانکار موظف است بلافاصله نسبت به رفع آن اقدام کند. و در صورت بروز مشکل در نحوه عملکرد سامانه مانند مشکل در اجرای سامانه به صورتی که به دلایلی کاملاً قابل استفاده نباشد و نیاز به راه اندازی مجدد داشته باشد، پیمانکار باید بلافاصله نسبت به حل مشکل به صورت تلفنی و حضوری در صورت نیاز اقدام کند. همچنین پیمانکار باید نسبت به، به روز رسانی سامانه و بازیابی اطلاعات در صورت بروز خطا، حمله یا سوء استفاده از سامانه، در مدت زمان پشتیبانی متعهد شود.

پیمانکار موظف است تا 24 ماه سامانه تحویل شده را به صورت رایگان پشتیبانی کند که شامل رفع خطاهای احتمالی و آموزش های موردی مورد نیاز کاربران می باشد و پس از آن تا 5 سال متعهد به پشتیبانی از آن می باشد.

بنابر این ضروری است پیمانکار شرایط و هزینه پشتیبانی را نیز در پروپوزال اعلام نماید.

نام پروژه: طراحی و پیاده سازی سامانه داده کاوی و متن کاوی بر روی داده های مرکز پژوهش و سنجش افکار صداوسیما

تاریخ آخرین تغییرات 30/10/96

سند: WEVR 950406

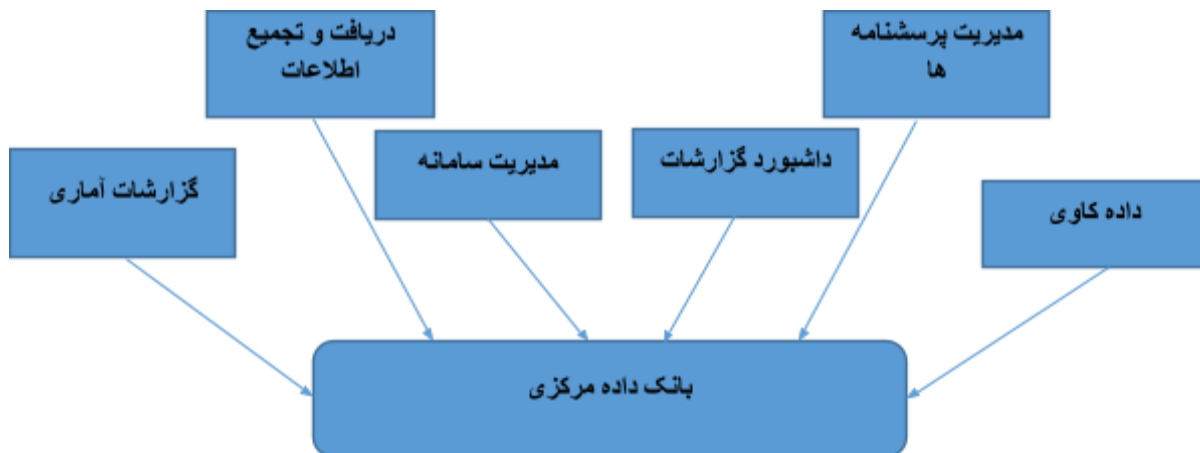
پیوست یک

مشخصات فنی
پروژه

پژوهش، طراحی و پیاده سازی سامانه نرم افزاری داده کاوی و متن کاوی
بر روی داده های مرکز پژوهش و سنجش افکار صداوسیما
باقابلیت پیش بینی رفتار مخاطبان

1- نیازهای عملکردی³ و امکانات و قابلیت های سامانه

امکانات و قابلیت هایی که باید در سامانه موردنظر دیده شود در شکل 1 نمایش داده شده است، و در بخش های بعدی توضیحات هر یک ارائه خواهد شد:



شکل 1- نمونه ی پیشنهادی از ورودی و زیرسیستم های سامانه

1-1 مدیریت سامانه

در قسمت مدیریت سامانه باید کاربران و سطوح دسترسی آن ها تعریف شود و در داشبورد مدیریتی امکان نمایش و ویرایش با توجه به سطوح دسترسی دیده شود. لازم است در بخش مدیریت سامانه موارد زیر دیده شود:

1. تعریف و ویرایش کاربران و سطوح دسترسی
2. باید دسترسی به این سامانه برای سطوح مختلف کاربران از اپراتور ورود اطلاعات تا مدیران میانی و ارشد تعریف شود.
3. ایجاد سطوح دسترسی جهت مشاهده فرم های گزارش گیری که متناسب با نیاز و اهمیت اطلاعات منتج شده از آن گزارش اختصاص داده می شود.
4. ایجاد داشبورد مدیریتی جهت مشاهده فعالیت های سامانه
5. رعایت اصل سهولت استفاده برای مدیران
6. مشاهده و مدیریت log های فعالیت های انجام شده در سامانه
7. امکان انجام تهیه نسخه پشتیبان از سامانه
8. امکان تنظیم و برقراری ارتباط با بانک های داده SQL Server توسط ادمین سامانه (تا کاربر بتواند بعد از برقراری این ارتباط با دیتابیس، Stored Prosedure مورد نظر خود را انتخاب کند و آن را اجرا نماید. Stored Prosedure ها از قبل در بانک اطلاعاتی تعریف شده اند).

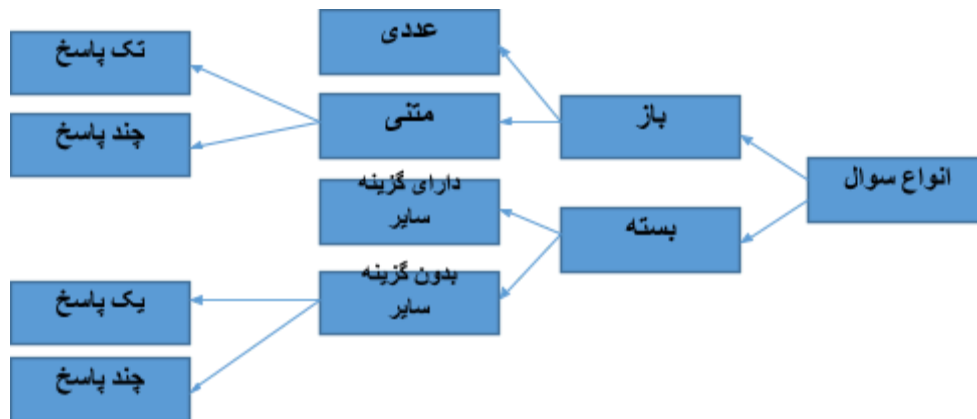
1-2 دریافت و تجمیع اطلاعات

³ Functional

در این بخش ورود اطلاعات نظرسنجی ها توسط اپراتور با استفاده از فرم ورود اطلاعات انجام می گیرد به این صورت که فایل نتایج نظرسنجی در سامانه import می شود و توسط کاربر، عنوان سؤالات اضافه می شود، و داده ها در بانک اطلاعاتی تجمع می شود. داده های ورودی در فرمت فایل های sav. (فایل ها csv. ، spss و خروجی های Stored Procedure از بانک اطلاعاتی SQLServer می باشد. اگر در فایل های ورودی sav، عناوین سؤالات، نوع، محدوده، مقدار و عناوین گزینه ها در قالب lable قرار داده شده باشد، برای سهولت ورود اطلاعات و کاهش کار اپراتور باید از موارد ذکر شده همان فایل خوانده شود و در فرم ورود اطلاعات که عناوین سؤالات نشان داده می شود باید قابلیت ویرایش موجود باشد. همچنین کاربر باید بتواند اطلاعات اضافی دیگر (متادیتا) مانند جامعه آماری (به صورت متنی)، تاریخ اجرا، شیوه اجرا (تلفنی، میدانی، اینترنتی) را برای پرسشنامه ها وارد کند که در ادامه جهت دسته بندی و گزارش گیری از این داده های افزوده استفاده شود.

کاربر باید بتواند Stored Procedure خود را از بانک اطلاعاتی به عنوان ورودی انتخاب نماید. پس از ورود موفق اطلاعات، کاربر به بخش مدیریت پرسشنامه ها هدایت شود و در غیر این صورت خطا به صورت مشخص و ذکر محل و جزئیات خطا نشان داده شود.

در پرسشنامه ها، انواع سؤالات در شکل 2 نشان داده شده است:



شکل 2- انواع سؤالات نظرسنجی ها

سوالات باز یا تشریحی به سؤالاتی گفته می شود که برای پاسخ دهنده گزینه مشخص نمی شود و می تواند متنی یا عددی باشد. چنانچه متن باشد می تواند یک و یا بیشتر پاسخ داشته باشد مانند انتقادات و پیشنهادات. سوالات بسته به سؤالاتی گفته می شود که برای پاسخ دهنده گزینه مشخص می شود، در این نوع سؤالات ممکن است یک قسمت باز هم داشته باشیم، برای مثال اگر بیننده برنامه ای را نگاه نمی کند دلایل آن می تواند بین چهار گزینه مشخص شده باشد در غیر این صورت گزینه سایر را انتخاب می کند که دلیل خود را به صورت متن در ادامه سؤال به صورت یک یا چند پاسخ وارد می کند.

1-3 مدیریت پرسشنامه ها

در این بخش موارد مرتبط با مدیریت پرسشنامه ها باید پیاده سازی شود.

1. مشاهده لیست پرسشنامه ها، با امکان جستجو و فیلتر بر اساس ویژگی هایی مانند عنوان، تاریخ ایجاد، تاریخ قفل و متادیتا
2. مشاهده تعداد کل پاسخ ها و آمار کلی در مورد هر پاسخنامه (تعداد این موارد باید به صورت پویا قابل اضافه شدن باشد).
3. قابلیت مشاهده و ویرایش اطلاعات اضافی دیگر (متادیتا) مانند جامعه آماری (به صورت متنی)، تاریخ اجرا، شیوه اجرا (تلفنی، میدانی، اینترنتی) نیز موجود باشد که در ادامه جهت دسته بندی و گزارش گیری از این داده های افزوده استفاده شود.
4. دسته بندی موضوعی و ویرایش پرسشنامه ها
5. امکان حذف، ویرایش و قفل کردن پرسشنامه ها
6. امکان ارتباط دهی نظرسنجی های مختلف و یکسان سازی پارامترهای مشابه ولی بانام های مختلف در نظرسنجی ها

1-4 گزارشات آماری

در این بخش کلیه گزارشات آماری باید اجرا شود که به شرح زیر می باشد:

1. امکان انتخاب و جستجوی یک یا چند پرسشنامه ها جهت انجام گزارشات آماری (در صورتی که ارتباط قبلاً بین پرسشنامه ها تعریف شده باشد)
2. جدول های فراوانی تکتک پرسش ها که شامل فراوانی، درصد و درصد تراکمی باشد.
3. جدول های متقاطع بین پرسش ها (حداکثر 5 مورد) با امکان انتخاب فراوانی خانه ها، درصد کل، درصد سطری، درصد ستونی و درصد های حاشیه ای.
4. امکان محاسبه مشخص های توصیفی (میانگین، میانگین پیراسته، انحراف معیار، دامنه، مقدار کمینه، مقدار بیشینه، میانه، نما، چندک ها، چولگی و کشیدگی)
5. امکان محاسبه نمره کل از مجموعه ای از پرسش ها و تعیین دامنه نمره کل.
- 5.1. نمره کل به صورت وزنی یا غیر وزنی باشد (در صورتی که وزن نداشتن سؤال، وزن یک به عنوان پیش فرض در نظر گرفته می شود). کاربر باید بتواند برای سؤالات وزن مورد نظر خود را تعریف نماید.
- 5.2. شیوه محاسبه شاخص انتخابی باشد (میانگین یا جمع)
- 5.3. بازه ی محاسبه شاخص انتخابی باشد.
6. امکان محاسبه ضرایب همبستگی پیرسون، اسپیرمن و تاو کندال و آزمون های آماری آن ها.
7. امکان محاسبه ضرایب وی کرامر، دی سامرز و ضریب توافقی و آزمون های آماری آن ها.
8. امکان محاسبه شاخص های ایتا دو و امگا دو.
9. امکان نمایش جدول و رسم نمودارهای دایره ای، میله ای، هیستوگرام، پراکنش و خطی با امکان چاپ (برای تمام موارد مندرج در این بخش، همچنین پارامتر نمره کل نیز در تمام موارد به عنوان یکی از گزینه ها در محاسبات قرار داده شود).
10. امکان اجرای آزمون های آماری خی دو، تی، اف (تجزیه واریانس)، کروسکال والیس، من ویتنی، میانه.
11. امکان اجرای آزمون گردش (تصادفی بودن نمونه، Run Test)
12. امکان محاسبه واریانس درون بلوکی و بین بلوکی و همبستگی درون خوشه ای (بلوکی)
13. امکان اجرای تمام محاسبات ذکر شده در زیر گروه ها (مثلاً محاسبه نمره کل برای گروه مردها و زن ها به طور مجزا)

14. ارائه خروجی های فارسی (بهویژه ارقام، درصدها و نتایج آزمون ها) باقابلیت استفاده در مجموعه Excel و Word.
15. با امکان تعریف پرسش تفکیک کننده جهت تفکیک گزارش های آماری
16. امکان ارائه تمام روش های سنجش همبستگی از قبیل:
 - 16.1. مدل معادلات ساختاری
 - 16.2. تحلیل عاملی (تحلیلی و اکتشافی) و آزمون های مرتبط شامل آزمون کرویت و KMO (توسط برنامه factor loading را نمایش دهد)
 - 16.3. محاسبه ضریب پایایی با انتخاب مجموعه ای از سوالات (همه انواع) به انتخاب کاربر
 - 16.4. محاسبه ضریب همبستگی چند رشته ای و چند رشته ای نقطه ای با انتخاب مجموعه ای از سوالات به انتخاب کاربر
17. آزمون نرمالیتی (انتخاب یک سؤال و بررسی نرمال بودن مقادیر توزیع شده)
18. امکان Clustering بر اساس ویژگی های انتخابی (بمصورت اکتشافی یا غیر اکتشافی) بر اساس سوالات تعیین شده cluster انجام دهد.
19. امکان export پاسخنامه هایی که باهم تجمیع شده در قالب سوالات ادغام شده به صورت که پرسشنامه هر رکورد در فایل تجمعی مشخص شده باشد.
20. امکان استفاده از روش های زیر برای جانهی (imputation) افراد دارای پرسش بی پاسخ به روش های:
 - نزدیکترین همسایه،
 - جانشینی پرسش همبسته،
 - جانشینی با انتخاب تصادفی،
 - جانشینی با انتخاب تصادفی،
 - جانهی رگرسیونی، جانهی نسبتی، و
 - جانهی میانگین به سه شکل میانگین فرد، میانگین پرسش و میانگین پرسش تصحیح شده

1-5 دادمکاو

در بخش دادمکاو داده های نظرسنجی ها که شامل موارد زیر است:

1. بررسی همبستگی میان پاسخ های سوالات مختلف در یک پرسشنامه و یا در چند پرسشنامه مرتبط به هم در این سیستم همبستگی (اعم از اسپیرمن و پیرسان) بین پاسخ های سوالات محاسبه می شود. چنانچه پرسشنامه های مختلفی از به عنوان مثال یک منطقه جغرافیایی جمع آوری شده باشد می توان این همبستگی را بین پاسخ پرسش های چند پرسشنامه را هم محاسبه و گزارش نمود.
- برقراری ارتباط بین پرسشنامه های مختلف نیز یکی از موارد نسبتاً چالشی این پروژه است. ارتباط بین نظرات در پرسشنامه هایی که در زمانهای مختلف در سازمان تدوین می شود در صورتی می تواند تحلیل گردند که برای موارد مشابه در زمانهای مختلف از نام یکسانی استفاده شده باشد. مثلاً همیشه از لغت سن استفاده شود یا همیشه از تاریخ تولد و نه اینکه در یک پرسشنامه سن باشد و در دیگری تاریخ تولد. تنها در صورت وجود موارد مشابه همنام امکان تحلیل ترکیبی نتایج چند پرسشنامه وجود دارد.
- چنانچه در یک محدوده زمانی مشخص دو نظرسنجی حول دو موضوع مختلف انجام شده باشد که یک موضوع بر دیگری اثرگذار باشد (به عنوان مثال نظرسنجی در مورد پخش یک برنامه سیاسی از تلویزیون و نظرسنجی در مورد انتخابات در یک بازه زمانی مثلاً یک هفته مانده به انتخابات) انتظار می رود بین این دو نظرسنجی با توجه به

ویژگی های مشابه از قبیل تحصیلات، سن، محدوده جغرافیایی و تمام موارد مشابه دیگر (ارتباط برقرار کرد و نتیجه ای یک موضوع را بر موضوع دیگر به دست آورد. (برای مثال بعد از پخش مناظره سیاسی کاندید X، افراد محدوده جغرافیایی که همزمان در هر دو نظرسنجی شرکت کرده اند یا محدوده سنی مشخص که در هر دو نظرسنجی شرکت کرده اند، چه رفتاری نسبت به کاندیدها پیدا کرده اند.)

2. ارائه الگوهای تکرار شونده پنهان در نتایج نظرسنجی ها
الگوهای پرتکرار یعنی ترکیبات چندتایی از پاسخ ها که درصد بالایی از پرسشنامه ها وجود دارد را گزارش می کند. به عنوان مثال اگر سہتایی (مرد، بیکار، نام یک کاندیدای خاص مثل آقای X) اگر در نسبت زیادی از پرسشنامه ها باشد بیانگر یک الگوی تکرار شونده است که بیانگر آن است که مردان بیکار به آقای X اقبال نشان داده اند. البته برای این قانون علاوه بر فراوانی، ضریب اطمینان هم حساب می شود که مطمئن باشیم قانون صحیحی استخراج کرده ایم.

3. یافتن افراد مشابه هم در جامعه و بیان مهم ترین ویژگی های آن ها
افراد بر اساس نحوه پاسخ به پرسش های مختلف با استفاده از الگوریتم های خوشه بندی گرومندی می گردد و سپس ویژگی های هر گروه به لحاظ آماری مشخص می گردد. پس از بررسی نهایی حداقل 3 الگوریتم خوشه بندی مناسب انتخاب می شود که به عنوان الگوریتم های خوشه بندی سیستم استفاده شود.

4. پیش بینی رفتار افراد بر اساس الگوهای استخراج شده
پس از خوشه بندی و گرومندی افراد این امکان وجود دارد که گروه یک فرد بر اساس پاسخ های مختلف مشخص شود و پاسخ پرسشی که فرد به آن جواب نداده است می تواند بر اساس پاسخ های گروه پیش بینی شود و خروجی برنامه پیش بینی پاسخ افراد بر اساس ویژگی های انتخابی داده شده ایشان خواهد بود.

5. در نهایت سامانه باید بتواند قوانین و روابط مختلف را با دقت حداقل 90% پیش بینی نماید. همچنین سامانه باید برای هریک از موارد همبستگی یک ضریب اطمینان و نرخ مثبت کاذب کمتر از 0.1 نیز گزارش کند تا کاربر درک بهتری از میزان ارزش هریک از روابط داشته باشد. همچنین امکان تنظیم دقت گزارشات توسط کاربر فراهم باشد (به عنوان مثال گزارشات بالای 70% را نشان دهد.).

6. قابلیت ارائه خروجی های فارسی گزارشات دادمکاو سامانه با قابلیت استفاده در مجموعه Excel و Word.

7. امکان تولید قالب گزارشی جهت ارائه در داشبورد مدیریتی

1-6 متن کاوی

در بخش متن کاوی داده های نظرسنجی ها که شامل موارد زیر است:

1. استخراج موضوعات اصلی نظرات متنی
سیستم کل متن را می گیرد و موضوعات کلیدی مهم را از آن استخراج می کند. برای هر موضوع بین 2 تا 10 کلمه به صورت انتخابی توسط کاربر برای مشخص شدن موضوع اضافه می کند. مثلاً موضوع زمان برنامه ها کلماتی مانند وقت، مدت، طول برنامه و موارد مشابه را برای تبیین موضوع اضافه می کند.

2. تحلیل افکار مطرح حول هر موضوع در پاسخ های پرسش های آزاد از لحاظ مثبت یا منفی بودن

در این بخش پس از استخراج موضوعات کلیدی سیستم با تحلیل نظرات مثبت و منفی حول هر موضوع علاوه بر کلمات کلیدی که مبین موضوع هستند یک امتیاز مثبت و یک امتیاز منفی نیز به آن موضوع اختصاص می‌دهد. این کار باعث می‌شود بجای اینکه از کل نظرات موجود در یک نظرسنجی، فقط مثبت یا منفی بودن نظرات کلی مشخص شود، به تفکیک موضوع، مثبت یا منفی بودن به همراه مقدار موارد مشخص شود.

3. یافتن علایق و نظرات گروه‌های مختلف شرکت‌کننده در نظرسنجی بر مبنای تفکیک رده سنی، جنسیت، میزان تحصیلات و محل سکونت و کل ویژگی‌های موجود در نظرسنجی‌ها.
- خروجی این بند نیز مشابه دو بند فوق است فقط کاربر می‌تواند در این بخش یکی از ویژگی‌های داده‌ها را مشخص کند و متون باز را بر اساس آن ویژگی فیلتر نماید. به عنوان مثال با مشخص کردن ویژگی مربوط به میزان تحصیلات و مشخص کردن مقدار این ویژگی مثلاً تحصیلات در سطح کارشناسی تمام تحلیل‌های بند 5 و 6 تنها برای متون جمع‌آوری شده از افراد با میزان تحصیلات کارشناسی انجام می‌پذیرد. به‌طور کلی این امکان در سیستم وجود خواهد داشت که بتوان افراد را بر مبنای یک ویژگی خاص انتخاب نماید و برای آن زیرمجموعه از افراد بررسی آماری روی پاسخ‌ها (که بیانگر نظرات و علایق افراد است) صورت می‌گیرد.
4. لازم است سناریوی تست و روایی سنجی و دقت گزارشات متن‌کاوی به تایید کارفرما برسد.
5. امکان تولید قالب گزارش‌های جهت ارائه در داشبورد مدیریتی

1-7 داشبورد گزارشات

این داشبورد به صورت یک تب در همان فهرست اصلی است نه یک داشبورد دیگر که احتیاج به نام کاربری دیگری دارد، فقط برنامه نویسی باید طوری باشد که چنانچه مدیر ارشد لاگین کرد بقیه تب‌های فهرست را نبیند و مستقیماً وارد صفحه داشبورد گزارشات شود. داشبورد مدیریتی باید قابلیت‌های زیر را دارا باشد:

1. قابلیت نمایش تمام گزارشات ایجاد شده در بخش‌های قبل با توجه به نوع دسترسی
2. قابلیت شخصی سازی داشبورد توسط کاربر نهایی فقط با چند کلیک
3. دسترسی به اطلاعات اصلی پرشنامه‌ها
4. ساده بودن و درک سریع بصری
5. قابلیت طراحی انواع نمودارهای متنوع آماری Pie ، میله ای و..
6. قابلیت نمایش گزارشات به صورت جدولی و نموداری
7. قابلیت تغییر رنگ و فونت و اندازه گزارش با سلیقه کاربر
8. قابلیت نمایش در موبایل

1-8 قابلیت‌های عمومی

1. تمامی سرویس‌ها باید مستقل از اینترنت و در شبکه اینترنت مرکز پژوهش قابلیت اجرا داشته باشند.
2. واسط کاربری تحت وب به صورت Responsive
3. جذابیت ابزار و محیط کاربرپسند

2- مشخصات سخت افزاری سرور

اعلام مشخصات سخت افزاری مورد نیاز برای سرور جهت اجرای سامانه تحت وب به عهده پیمانکار و با تأیید کارفرما می باشد.

پیوست دو

مستندات تحویل گیری نرم افزار
پروژه

پژوهش، طراحی و پیاده سازی سامانه نرم افزاری داده کاوی و متن کاوی
بر روی داده های مرکز پژوهش و سنجش افکار صداسیما
باقابلیت پیش بینی رفتار مخاطبان

1. فاز روشن کردن صورت مسئله

- 1.1. ارزیابی زیرساخت کنونی: در این فاز، محدوده پروژه، اهداف، توصیه ها، وضعیت سیستم فعلی، معماری شبکه، زیرساخت و سرویس کنونی و آنالیز وضعیت کاربردی سیستم جدید و آنالیز امنیتی بررسی می شود.
- 1.2. ساختار پروژه: در این بخش روش ها، اهداف، فرضیه ها، محدودیت های پروژه، همچنین محدوده پروژه، trade off و برآورد پروژه ذکر می شود.
- 1.3. گزارش حوزه چشم انداز⁴: در این بخش چشم انداز پروژه، فرصت های کسب و کار، تشریح راحل ارائه شده، اهداف، فرضیات و محدودیت ها، نیازمندی های کاربری، سیستمی و عملیاتی ذکر می شود. لیست فیچرها و عملکردها، مواردی که در محدوده پروژه می گنجد/نمی گنجد، ذکر می شود. استراتژی طراحی تکنیکی، استراتژی طراحی معماری و استراتژی طراحی راحل بیان می شود.

⁴ Vision scope

2. فاز برنامه ریزی

2.1. برنامه ها

- 2.1.1. طرح دسترس پذیری: در این بخش وضعیت جاری سیستم نظیر فرایندهای موجود، پرسشنامه های موجود و نقش های موجود بررسی می شود.
- 2.1.2. طرح Backup و Recovery: در این طرح شرحی از راحل، متدهای backup و recovery و گام های recovery مدنظر عنوان می شود.
- 2.1.3. طرح capacity: در این طرح، استراتژی های ارزیابی راحل کلی و کارایی مؤلفه ها و استفاده از این اطلاعات برای توسعه طرح استفاده، پیکربندی و بروز رسانی مؤلفه ها بیان می شود. در این طرح همچنین سناریوهای استفاده، ماتریس بار- وظیفه، انتظارات، ذخایر سیستم، ظرفیت مؤلفه ها، تنگناها، مانیتورینگ و متریک ها و راحل کلی ذکر می شود.
- 2.1.4. طرح ارتباطات: در این طرح قوانین ارتباطاتی که باید بکار گرفته شود و ماتریس سازمانی ارتباطات و جلسات بایستی ذکر شود. همچنین باید فرایندها، روش ها و ابزارهای مورد نیاز برای اطمینان از جمع آوری به موقع و مناسب، توزیع و مدیریت اطلاعات پروژه برای همه شرکت کنندگان در پروژه شناسایی شود.
- 2.1.5. طرح استقرار: در این طرح فاکتورهای لازم برای یک استقرار ملایم و انتقال به عملیات گذار توصیف می شود. این موضوع، شامل فرایندهای آمادسازی، نصب، آموزش، تثبیت و انتقال می باشد که شامل جزئیاتی درباره سناریوهای نصب و راه اندازی (به عنوان مثال در ادارات راه دور یا شعب)، مانیتورینگ استقرار و تایید راحل جدید است.
- 2.1.6. طرح توسعه: در این گزارش، فرایند توسعه راحل مورد استفاده برای پروژه باید توصیف شود. این سند مشخصات عملکردی که جزئیات فنی را پوشش می دهد، را در بر می گیرد.
- 2.1.7. طرح پشتیبانی کاربر نهایی: این طرح نحوه پشتیبانی از کاربر نهایی را در بر می گیرد و شامل نیازمندی های مورد استفاده نظیر ارتباطات فنی، آموزش کاربر نهایی، طرح گرافیکی و ... می باشد.
- 2.1.8. طرح جامع پروژه: در این طرح باید مواردی نظیر ساختار شکست کار، برنامه هر فرد، طرح توسعه، طرح آزمون، طرح ارتباطات، طرح پشتیبانی میکروسافت، طرح بودجه، طرح آموزش و ... پوشش داده شود.
- 2.1.9. طرح پشتیبانی: در این طرح باید مواردی نظیر حوزه پشتیبانی، مدل مرجع پشتیبانی، فعالیت های پشتیبانی نظیر گزارش وضعیت و گزارش مشکل باید لحاظ شود.
- 2.1.10. طرح مانیتورینگ: در این طرح باید پیش بینی نقاط شکست، مانیتورینگ آستانه منابع، تشخیص شکست، بازیابی شکست لحاظ شود.
- 2.1.11. طرح عملیات: در این طرح باید زیرساخت های عملیات، مدل گزارش دهی، نیازمندی های مهارتی، فرایند تغییر مدیریت و فعالیت های عملیاتی (راحل محصول/ویژگی) ذکر شود.
- 2.1.12. طرح کارایی: در این طرح باید نیازمندی های کارایی (زمان پاسخگویی)، محیط کارایی (زیرساخت، ابزارها و متدولوژی، مراحل آزمون عملکردی) ذکر شود.
- 2.1.13. طرح پایلوت: در این طرح باید شرح پایلوت (شامل معیار انتخاب پایلوت، حوزه، مشخصات کاربر هدف)، آمادسازی کاربر (طرح آموزش کاربر)، طرح انتقال (استراتژی انتقال، اضافه کردن کاربر بیشتر، عقبگرد، لغو)، ارزیابی پایلوت ذکر گردد.

2.1.14. طرح امنیت: در این طرح باید توصیف راحل (تخصیص نقش های امنیتی، توصیف کلی راحل، محیط اجرای راحل)، تهدیدات، کنترل های مدیریت، کنترل های عملوندی (کنترل های کارکنان، سطوح حساسیت، محدودیت دسترسی کاربر؛ تفکیک وظایف)، حفاظت فیزیکی و محیطی، کنترل های تکنیکی ذکر گردد.

2.1.15. طرح پشتیبانی: در این طرح باید منابع پشتیبانی (توافق نامه سطح خدمات، help desk (شامل log گیری، روال ها، آموزش، صدور گواهی نامه، تضمین کیفیت)، نیروی انسانی)، فرایند پشتیبانی، بهبود پشتیبانی، ذکر گردد.

2.1.16. طرح آزمون: در این طرح باید مفروضات و روش آزمون، میزان پاسخگویی آزمون، ویژگی ها و عملکردهایی که باید تست شود، نتایج مورد انتظار از آزمون، تحویل دانی های آزمون (مستندات و داده های آزمون)، روال های آزمون (روش ها و تنظیمات آزمون)، پیگیری و گزارش وضعیت، متدها و ابزارهای گزارش باگ ها ذکر گردد.

2.1.17. طرح آموزش: در این طرح باید نیازمندی های آموزش (مدیریت محصول، مدیریت برنامه، توسعه، آزمون، تجربه کاربری)، برنامه زمان بندی، روش آموزش، منابع و نیازمندی های آموزش ذکر گردد.

2.2. مشخصات

2.2.1. نیازمندی های کسب و کار: در این طرح باید خلاصه نیازمندی های کسب و کار، تحلیل هزینه- سودمندی، مقیاس پذیری برای نیازمندی های کسب و کار، امنیت قوی، پشتیبانی از انواع دستگاه ها، نیازمندی های کارایی، پاسخگویی و حسابرسی برگشت به حامیان کسب و کار ذکر گردد.

2.2.2. مشخصات عملکردی: در این طرح باید خلاصه ای از چشم انداز / محدوده پروژه، تاریخچه پروژه، خلاصه اجرایی مشخصات عملکردی، توجیه پروژه و اهداف طراحی (خلاصه نیازمندی های کسب و کار، خلاصه نیازمندی های کاربر، خلاصه نیازمندی های سیستم، خلاصه نیازمندی های عملیاتی، خلاصه سناریوهای استفاده و use case ها)، سناریوهای پشتیبانی نشده، مفروضات و وابستگی ها، طراحی راحل ذکر گردد.

2.2.3. طرح منطقی: در این طرح باید خلاصه طراحی منطقی، اشیاء، رفتارها، Attribute ها، روابط ذکر گردد.

2.2.4. طرح مفهومی: در این طرح باید خلاصه طرح مفهومی، وضعیت کسب و کار، بررسی اجمالی راحل مفهومی پیشنهادی، معماری راحل ذکر گردد.

2.2.5. نیازمندی های عملیاتی: در این طرح باید خلاصه نیازمندی های عملیاتی، مقیاس پذیری، امنیت، قابلیت مدیریت، قابلیت پشتیبانی، نیازمندی های نیروی انسانی، نیازمندی های آموزش نیروی انسانی و توافق سطح سرویس و نیازمندی های 7*24*365 ذکر گردد.

2.2.6. طرح فیزیکی: در این طرح باید خلاصه طراحی فیزیکی، محدودیت های محیطی و مفروضات، وابستگی ها (وابستگی های پروژه، وابستگی های محیط سخت افزار، وابستگی های محیط نرم افزار)، توسعه برنامه کاربردی (سرویس های کاربر (UI) و کامپوننت ها، سرویس های داده، توسعه زیرساخت (توپولوژی راحل، ارتباطات شبکه ای)، استراتژی امنیت، نصب و راه اندازی (نیازمندی سخت افزاری و نرم افزاری)، سرویس های کسب و کار، نصب UI ذکر گردد.

2.2.7. نیازمندی های سیستم: در این طرح باید خلاصه نیازمندی سیستم، وابستگی های سیستم ها و سرویس ها، نیازمندی های قابلیت همکاری، ارزیابی اثرات شبکه و تأثیر نسخه گذاری روی سیستم و محیط ذکر گردد.

2.2.8. سناریوی استفاده: در این طرح باید سناریوهای استفاده (2 , use case1 , ...) ذکر گردد.

2.2.9. نیازمندی های کاربر: در این طرح باید تجربه کاربری، سهولت استفاده، قابلیت اطمینان، کارایی، نیازمندی های چند زبانی، دسترسی و نیازمندی های آموزش کاربر نهایی ذکر گردد.

3. فاز توسعه دادن

- 3.1. گزارش آزمون و باگ: در این طرح باید خلاصه گزارش آزمون و باگ ها، روش ها و ابزارهای آزمون، گزارش آزمون، محدوده آزمون، گزارش باگ ها، طرح تجدیدنظر در آزمون ها ذکر گردد.
- 3.2. بررسی قابلیت ردیابی⁵: در این طرح باید ممیزی محدوده و ابعاد آن، ممیزی مشخصات کاربردی، ممیزی طرح جامع پروژه، ممیزی طرح پشتیبانی کاربر نهایی، ممیزی طرح توسعه، ممیزی طرح آموزش، ممیزی طرح مهاجرت، ممیزی طرح ابزارها ذکر گردد.

4. فاز ثبات و پایداری

- 4.1. بازبینی پایلوت: در این طرح باید فاکتورها و معیارهای موفق، خلاصه فیدبک فاز پایلوت، فیدبک تیم پروژه و نتایج و توصیه های فاز پایلوت ذکر گردد.
- 4.2. مشخصات آزمون: در این طرح باید خلاصه مشخصات آزمون، مشخصات ورودی، مشخصات خروجی، محیط آزمون، نیازمندی های رویه های ویژه، نیازمندی های پیش از آزمون، موارد آزمون ها (فیچرها) ذکر گردد.

5. فاز استقرار

- 5.1. Post Project Analysis: در این طرح باید برای موارد زیر، دستاوردها، چالش ها، تجربیات کسب شده ارائه گردد. لیست موارد: برنامه ریزی، منابع، مدیریت پروژه و زمان بندی، توسعه/ طراحی/ مشخصات، آزمون، ارتباطات، تیم/ سازمان، راحل و ابزارها.
- 5.2. گزارش خاتمه پروژه: در این گزارش باید خلاصه گزارش خاتمه پروژه، تحویل چشم انداز، تغییراتی که در چشم انداز نهفته است (تغییرات در کسب و کار سازمان، تغییرات در چشم انداز، تغییرات در تیم، تغییرات در فرایند مشتری، تغییرات در فرایند پروژه، تغییرات در طرح، تغییرات در مشخصات، تغییرات در جدول زمانی) ذکر گردد.

⁵ Traceability Audit