

Phần 2. YÊU CẦU VỀ KỸ THUẬT
Chương V. YÊU CẦU VỀ KỸ THUẬT

1. Giới thiệu chung về dự án/dự toán mua sắm, gói thầu:

- Tên gói thầu: Số hóa tài liệu lưu trữ của Văn phòng Đảng ủy Đặc Khu Vân Đồn
- Chủ đầu tư: Văn phòng Đảng ủy Đặc Khu Vân Đồn
- Nguồn vốn: Vốn ngân sách tỉnh cấp bổ sung có mục tiêu năm 2026.
- Thời gian thực hiện gói thầu: 30 ngày.
- Địa điểm cung cấp dịch vụ: Văn phòng Đảng ủy Đặc Khu Vân Đồn.
- Nội dung phạm vi, quy mô:

STT	Danh mục dịch vụ	Khối lượng mời thầu	Đơn vị tính	Địa điểm thực hiện dịch vụ	Ngày hoàn thành dịch vụ
(1)	(2)	(4)	(5)	(6)	(7)
1	Quét tài liệu	250.000	Trang A4 quy đổi	Văn phòng Đảng ủy Đặc Khu Vân Đồn	30 ngày
2	Nhập các trường dữ liệu có số lượng ký tự trong 1 trường (n) ≤ 15	1.333.333	Trường	Văn phòng Đảng ủy Đặc Khu Vân Đồn	30 ngày
3	Nhập các trường dữ liệu có số lượng ký tự trong 1 trường 15 < (n) ≤ 50	250.000	Trường	Văn phòng Đảng ủy Đặc Khu Vân Đồn	30 ngày
4	Nhập các trường dữ liệu có số lượng ký tự trong 1 trường (n) > 50	166.667	Trường	Văn phòng Đảng ủy Đặc Khu Vân Đồn	30 ngày
5	Ổ cứng lưu trữ - 8TB HDD	02	Cái	Văn phòng Đảng ủy Đặc Khu Vân Đồn	30 ngày

2. Mục tiêu công việc:

Thực hiện Số hóa tài liệu lưu trữ của Đảng ủy Đặc khu Vân Đồn để xây dựng kho cơ sở dữ liệu điện tử tài liệu lưu trữ; từng bước chuyển phương thức hoạt động của lưu trữ truyền thống sang lưu trữ điện tử; chuyển đổi việc khai thác tài liệu lưu trữ trực tiếp bản gốc bằng giấy sang khai thác dưới dạng file điện tử, giúp cho việc khai thác cung cấp thông tin được nhanh chóng, hiệu quả và tiết kiệm thời gian; góp phần cải cách hành chính, nâng cao hoạt động ứng dụng công nghệ thông tin trong công tác lưu trữ, kéo dài tuổi thọ của văn bản gốc (bằng giấy) đảm bảo có các bản sao lưu dự phòng tài liệu lưu trữ gốc và hỗ trợ việc thực hiện duy tu, bảo dưỡng, kiểm soát tài liệu.

3. Yêu cầu kỹ thuật của gói thầu:

3.1. Yêu cầu thực hiện

- Thực hiện công tác chỉnh lý, số hóa các hồ sơ, tài liệu theo hướng hiện đại hóa, đúng tiêu chuẩn theo quy định của Nhà nước, phục vụ quản lý có hiệu quả các quy trình nghiệp vụ công tác lưu trữ tại Đảng ủy Đặc khu Vân Đồn, từng bước đưa công tác lưu trữ tài liệu được chuẩn mực, hiệu quả và thực hiện thống nhất, đúng với quy định của Nhà nước.

- Tài liệu lưu trữ của mỗi hồ sơ sau khi chỉnh lý hoàn tất sẽ tiếp tục thực hiện số hóa, lưu khi quét (scan) xong, được lưu dưới dạng các file điện tử và tích hợp vào cơ sở dữ liệu. Tập tin điện tử phải được định dạng theo những tiêu chuẩn phổ biến với siêu dữ liệu (metadata) có tính cấu trúc và kỹ thuật cơ bản để đảm bảo khả năng bảo quản tập tin lâu dài. Tối ưu hóa kích cỡ tập tin để giảm thiểu gánh nặng lưu trữ và lưu thông trên mạng;

- Quá trình số hóa tài liệu lưu trữ phải bảo đảm tính pháp lý và tính trung thực của hồ sơ; tuân thủ các quy trình kỹ thuật tại Hướng dẫn số 40-HD/VPTW ngày 07/11/2018 của Văn phòng Ban chấp hành Trung ương hướng dẫn số hóa tài liệu tại các cơ quan, tổ chức đảng, tổ chức chính trị - xã hội; Thông tư số 05/2025/TT-BNV ngày 14/5/2025 của Bộ Nội vụ về quy định nghiệp vụ lưu trữ tài liệu lưu trữ số; Công văn số 14748-CV/VPTW ngày 06/5/2025 của Văn phòng Trung ương Đảng về việc sửa đổi, bổ sung một số điểm của Hướng dẫn số 40-HD/VPTW, ngày 07/11/2018 về số hoá tài liệu.

3.2. Yêu cầu về tiêu chuẩn kỹ thuật áp dụng

Việc thực hiện triển khai đảm bảo tuân thủ danh mục tiêu chuẩn kỹ thuật công bố kèm theo Thông tư số 39/2017/TT-BTTTT ngày 15/12/2017 của Bộ Thông tin và Truyền thông. Cụ thể:

STT	Loại tiêu chuẩn	Ký hiệu tiêu chuẩn	Tên đầy đủ của tiêu chuẩn	Quy định áp dụng
2	Tiêu chuẩn về tích hợp dữ liệu			
2.1	Ngôn ngữ định dạng văn bản	XML v1.0 (5 th Edition)	Extensible Markup Language version 1.0 (5 th Edition)	Bắt buộc áp dụng một trong hai tiêu chuẩn
		XML v1.1 (2 nd Edition)	Extensible Markup Language version 1.1	
2.2	Ngôn ngữ định dạng văn bản cho giao dịch điện tử	ISO/TS 15000:2014	Electronic Business Extensible Markup Language (ebXML)	Bắt buộc áp dụng
2.3	Định nghĩa các lược đồ	XML Schema	XML Schema version	Bắt buộc áp dụng

STT	Loại tiêu chuẩn	Ký hiệu tiêu chuẩn	Tên đầy đủ của tiêu chuẩn	Quy định áp dụng
	trong tài liệu XML	V1.1	1.1	
2.7	Trình diễn bộ kí tự	UTF-8	8-bit Universal Character Set (UES)/Unicode Transformation Format	Bắt buộc áp dụng
3	Tiêu chuẩn về truy cập thông tin			
3.4	Văn bản	(.txt)	Định dạng Plain Text (.txt): Dành cho các tài liệu cơ bản không có cấu trúc	Bắt buộc áp dụng
		(.rtf) v1.8, v1.9.1	Định dạng Rich Text (.rtf) phiên bản 1.8, 1.9.1: Dành cho các tài liệu có thể trao đổi giữa các nền khác nhau	Bắt buộc áp dụng
		(.docx)	Định dạng văn bản Word mở rộng của Microsoft (.docx)	Khuyến nghị áp dụng
		(.pdf) v1.4, v1.5, v1.6, v1.7	Định dạng Portable Document (.pdf) phiên bản 1.4, 1.5, 1.6, 1.7: Dành cho các tài liệu chỉ đọc	Bắt buộc áp dụng một, hai hoặc cả ba tiêu chuẩn
		(.doc)	Định dạng văn bản Word của Microsoft (.doc)	
		(.odt) v1.2	Định dạng Open Document Text (.odt) phiên bản 1.2	
3.5	Bảng tính	(.csv)	Định dạng Comma separated Variable/Delimited (.csv): Dành cho các bảng tính cần trao đổi giữa các ứng dụng khác	Bắt buộc áp dụng

STT	Loại tiêu chuẩn	Ký hiệu tiêu chuẩn	Tên đầy đủ của tiêu chuẩn	Quy định áp dụng
			nhau.	
		(.xlsx)	Định dạng bảng tính Excel mở rộng của Microsoft (.xlsx)	Khuyến nghị áp dụng
		(.xls)	Định dạng bảng tính Excel của Microsoft (.xls)	Bắt buộc áp dụng một hoặc cả hai tiêu chuẩn
		(.ods) v1.2	Định dạng Open Document Spreadsheets (.ods) phiên bản 1.2	
3.7	Ảnh đồ họa	JPEG	Joint Photographic Expert Group (.jpg)	Bắt buộc áp dụng một, hai, ba hoặc cả bốn tiêu chuẩn
		GIF v89a	Graphic Interchange (.gif) version 89a	
		TIFF	Tag Image File (.tif)	
		PNG	Portable Network Graphics (.png)	
3.13	Bộ ký tự và mã hóa	ASCII	American Standard Code for Information Interchange	Bắt buộc áp dụng
3.14	Bộ ký tự và mã hóa cho tiếng Việt	TCVN 6909:2001	TCVN 6909:2001 “Công nghệ thông tin - Bộ mã ký tự tiếng Việt 16-bit”	Bắt buộc áp dụng
3.15	Nén dữ liệu	Zip	Zip (.zip)	Bắt buộc áp dụng một hoặc cả hai tiêu chuẩn
		.gz v4.3	GNU Zip (.gz) version 4.3	

3.3. Yêu cầu về số hóa tài liệu

3.3.1 Nguyên tắc số hóa hồ sơ

- Việc thực hiện quét (scan) phải tuân thủ theo nguyên tắc về nghiệp vụ lưu trữ và kỹ thuật công nghệ thông tin;

- Không số hoá tài liệu tuyệt mật. Đối với tài liệu tối mật, tài liệu mật, hồ sơ nhân sự phải trình cấp có thẩm quyền phê duyệt. Việc số hoá tài liệu bí mật nhà nước phải tuân thủ quy định của Đảng, pháp luật của Nhà nước về bảo vệ bí mật nhà nước và phải áp dụng giải pháp mã hoá, bảo mật của Ban Cơ yếu Chính phủ.

- Khi số hoá phải bảo đảm chính xác, giống nguyên trạng về nội dung, bố cục so với tài liệu gốc.

- Việc scan số hóa tài liệu lưu trữ phải được thực hiện tại địa điểm được chủ đầu tư bố trí để đảm bảo việc bảo mật và an toàn thông tin ở mức cao nhất;

- Phải đảm bảo tính chính xác, giữ nguyên bố cục của tài liệu gốc;

- Việc sử dụng các định dạng tệp văn bản điện tử sau khi quét phải thống nhất định dạng có đuôi .pdf, .tiff;

- Việc lưu trữ, sắp xếp các file điện tử phải được ký hiệu thống nhất, theo trình tự khoa học, an toàn thông tin, có khả năng truy cập, khai thác ngay khi tài liệu được tạo lập.

3.3.2 Yêu cầu về Kỹ thuật trong công tác thi công

- Đơn vị thi công cần triển khai đầy đủ thiết bị thi công đảm bảo triển khai công việc theo đúng tiến độ đề ra.

- Thuyết minh, trình bày chi tiết, đầy đủ các bước thực hiện Số hóa tài liệu khối Đảng theo quy định hiện hành.

- Đơn vị thi công sẵn sàng trình diễn giải pháp phần mềm thi công số hóa, phương án thi công khả thi (nếu Chủ đầu tư có yêu cầu).

3.3.3 Yêu cầu về Thiết bị quét

Để đảm bảo chất lượng thi công số hóa tài liệu, khi thực hiện thi công Đơn vị triển khai phải đảm bảo cung cấp đầy đủ trang thiết bị để phục vụ thực hiện nhiệm vụ. Đối với từng loại tài liệu hồ sơ sẽ thực hiện trên các loại máy quét khác nhau ví dụ:

- Tài liệu hồ sơ bản giấy rời lẻ hoặc có thể tháo rời sẽ tiến hành quét trên máy quét cuộn giấy tự động;

- Các tài liệu mỏng, giấy rách nát có khả năng hư hỏng khi đưa vào máy tự động hoặc tài liệu dạng quyển không thể tháo rời thì sẽ tiến hành quét trên máy quét phẳng;

- Tương thích với tất cả các chuẩn máy quét hiện nay, hỗ trợ các chuẩn ISIS, TWAIN và WIA;

- Cho phép quét theo lô (batch), mỗi lô ứng với một hồ sơ;

- Hỗ trợ tự động phân tách tài liệu bằng mã vạch (barcode);

- Cho phép thực hiện quét lại trang tài liệu trong trường hợp chất lượng ảnh quét không đảm bảo;

- Cho phép chèn thêm, xóa hay thay đổi thứ tự ảnh quét trong tài liệu;

- Cho phép thực hiện chỉnh sửa ảnh quét: xoay trang 90⁰, 180⁰, chỉnh thẳng ảnh nghiêng;

- Các thiết bị thi công phải đáp ứng tối thiểu về số lượng và yêu cầu như sau (Nhà thầu có thể đề xuất thiết bị cấu hình tương đương/ hoặc cao hơn):

Stt	Tên thiết bị và cấu hình	Số lượng	Ghi chú
1	Máy scan khổ A3 kèm quét phẳng khổ A3	02 cái	Scan tài liệu rời lẻ khổ A3, tài liệu mỏng, tài liệu cũ, tài liệu đóng quyển
	- Độ phân giải quang học: ≥ 600 dpi - Tốc độ quét tự động: ≥ 80 tờ/phút ở chế độ màu 300 dpi - Cổng kết nối: USB 3.2 thế hệ 1x1		
2	Máy scan cuộn giấy tự động khổ A4	10 cái	Scan tài liệu rời lẻ khổ A4, tài liệu mỏng, tài liệu cũ
	- Độ phân giải quang học: ≥ 600 dpi - Tốc độ quét: ≥ 80 tờ/phút ở chế độ màu 300 dpi - Cổng kết nối: USB 3.2 thế hệ 1x1		
3	Máy tính bộ	14 bộ	Máy bộ để thực hiện công tác số hóa và nhập liệu
	- Bộ vi xử lý: \geq Core i5 thế hệ 12 trở lên - RAM: ≥ 8 GB DDR4 - Ổ đĩa cứng: ≥ 250 GB SSD - Win11 Pro bản quyền - Màn hình: ≥ 21 inch		
4	Máy chủ bộ	01 bộ	Cài đặt, vận hành, sao lưu dữ liệu Phần mềm quản lý dây chuyền số hóa trong quá trình thi công.
	- Bộ vi xử lý: \geq Intel – Xeon – Gold 5416S, 32GB Dual Rank x8 DDR5, 96W Smart Stg Li-ion Batt 145mm Kit, MR408i-o Gen11 SPDM Storage Cntlr, BCM 5719 1Gb 4p BASE-TOCP Adptr, 1600W FS Plat Ht Plg LH Pwr Sply Kit P52534-B21.		
5	Thiết bị chuyển mạch	01 bộ	Kết nối, tạo mạng nội bộ cho các thiết bị trên.
	Switch 24 cổng 10/100/1000Mbps		
6	Máy in A4	02 Cái	

3.3.4 Yêu cầu phần mềm thi công số hóa

Các phần mềm sử dụng trong quá trình thi công phải đáp ứng các yêu cầu chức năng như sau (hoạt động trong môi trường mạng nội bộ, không kết nối internet) có giao diện người dùng dạng Windows Forms (Desktop application) để đảm bảo thuận tiện cho người sử dụng, khả năng đáp ứng được yêu cầu năng suất cao trong công việc, tiến độ triển khai. Phần mềm phải đáp ứng các yêu cầu chức năng tối thiểu như sau:

a. Phần mềm quản lý dây chuyền số hóa và biên mục tài liệu

- Phần mềm dây chuyền số hóa phải có chứng nhận đánh giá An toàn thông tin.
- Có giao diện cho phép định nghĩa các loại tài liệu cần biên mục. Định nghĩa các trường dữ liệu ứng với từng loại tài liệu;
 - Có thể tạo ra hệ thống danh mục và gán vào các trường tin để người dùng chọn trong quá trình biên mục mà không cần phải nhập liệu;
 - Cho phép tạo ra các dự án, thông tài liệu và mục lục tài liệu;

- Thiết lập thời gian bắt đầu dự án, kết thúc dự án, quản lý tiến độ dự án;
- Cho phép thiết lập hệ thống dữ liệu phù hợp với phương án chính lý và tài liệu thực tế trong kho;
- Cho phép người dùng chọn kiểu hiển thị màn hình nhập liệu;
- Cho phép quản lý danh mục các hồ sơ tài liệu cần số hóa, các hồ sơ tài liệu đã số hóa xong;
- Cho phép phân quyền cho các nhóm với các chức năng tương ứng cho từng nhóm. Và phân quyền cho từng đối tượng cụ thể phù hợp với công việc được giao;
- Cho phép quản lý lỗi và thông báo cho các bộ phận liên quan sửa lỗi;
- Cho phép nhận dạng lại một vùng ảnh hoặc cả trang của tài liệu.
- Người dùng có thể giao tiếp trao đổi với nhau qua phần mềm
- Người dùng có thể kiểm soát được tiến độ xử lý công việc của mình thông qua công cụ trực quan trên phần mềm.
- Tất cả dữ liệu đều được lưu tập trung vào phần mềm quản lý đây chuyên số hóa, dữ liệu sẽ được sao lưu hàng ngày;
- Có chức năng giao việc, kiểm tra lại kết quả giao việc cho nhân viên nhập liệu.
- Cho phép kiểm tra và xác nhận trạng thái đến từng trường dữ liệu chưa đạt yêu cầu.
- Cho phép chuyển những tài liệu chưa đạt yêu cầu đến người biên mục để hoàn thiện;
- Màn hình soát lỗi chia làm 3 vùng: vùng ảnh, vùng kết quả nhận dạng, vùng ảnh phóng to và con trỏ sẽ đồng bộ giữa 3 vùng màn hình cho phép người sử dụng nhanh chóng hiệu chỉnh kết quả nhận dạng;
- Phần mềm cần có các bộ hàm API cho phép tích hợp với Phần mềm quét, Phần mềm quản lý tài liệu.
- Hệ thống phần mềm thi công số hóa tài liệu: cho phép kiểm soát tự động quy trình thi công, quản lý người dùng, Quản lý tiến độ, chất lượng quét tài liệu và nhập liệu, kiểm tra soát lỗi; hỗ trợ lập báo cáo hiện trạng quá trình thi công theo yêu cầu. Hỗ trợ thống kê hiện trạng của quy trình thi công.
- Hệ thống phần mềm thi công số hóa có áp dụng công nghệ nhận dạng tiếng Việt và phân loại, bóc tách thông tin tự động tài liệu tiếng Việt (do đơn vị tại Việt Nam phát triển để đảm bảo khả năng về thời gian triển khai và tính bảo mật của dữ liệu).
- Hệ thống phần mềm thi công số hóa có chức năng xuất ra định dạng pdf 2 lớp, định dạng word, excel, ... và phải giữ được cấu trúc văn bản như bản gốc.
- Hệ thống phần mềm thi công số hóa có áp dụng công nghệ nén dữ liệu để giảm dung lượng lưu trữ tài liệu lưu trữ, dung lượng không vượt quá 100kb đối với tài liệu khổ A4, chế độ trắng đen/xám, độ phân giải 300 dpi.

- Hệ thống phần mềm thi công số hóa tài liệu được cơ quan nhà nước đã cấp chứng nhận về bảo mật an toàn thông tin không có lỗ hổng.

b. Phần mềm số hóa tài liệu tiếng Việt

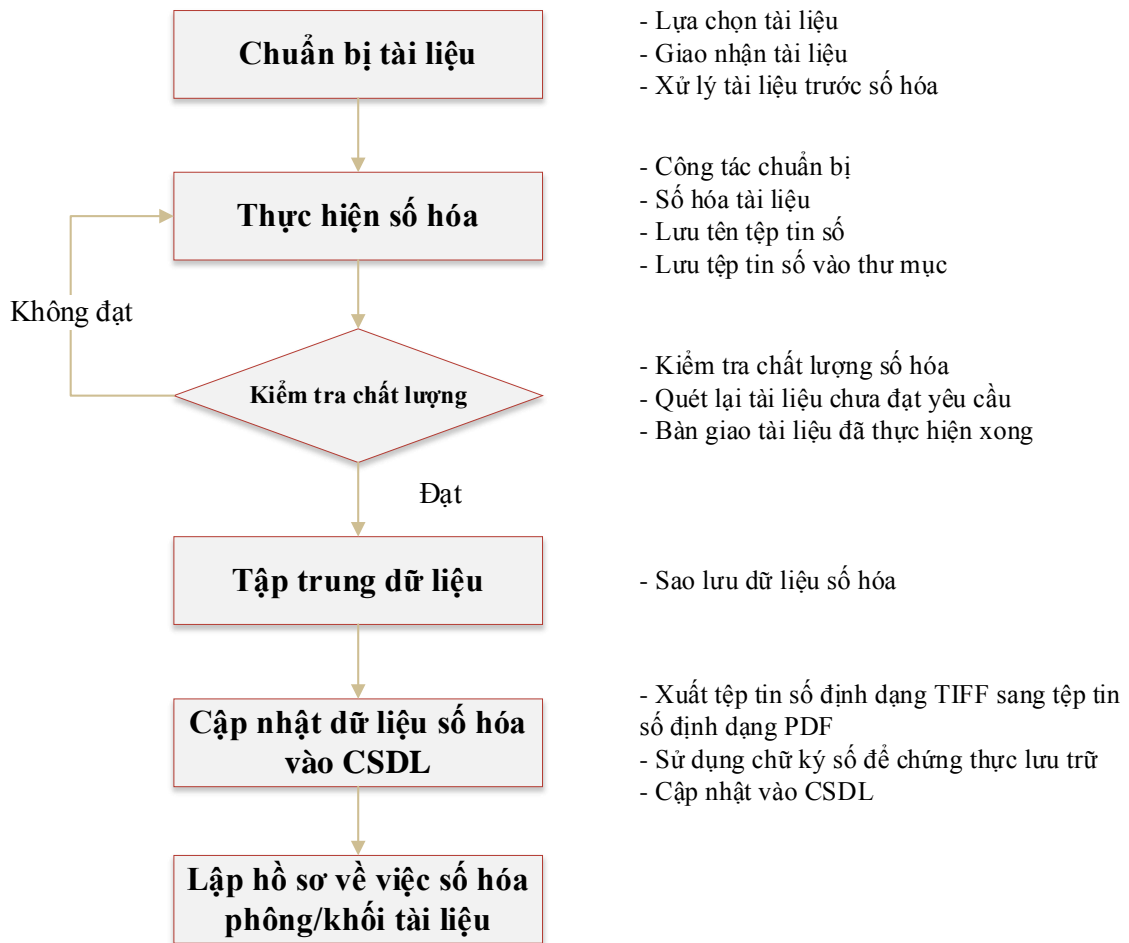
- Nhận dạng được nhiều ngôn ngữ (trong đó có tiếng Việt, tiếng Anh, tiếng Pháp, tiếng Nga) với độ chính xác mức ký tự tối thiểu là 97% đối với tài liệu mới;
- Nhận dạng tài liệu đa ngôn ngữ (tài liệu có đồng thời tiếng Việt và một trong những ngôn ngữ sau: Anh, Pháp, Nga)
 - Hỗ trợ kết xuất tài liệu số ra các định dạng PDF 2 lớp
 - + Lớp trên là ảnh của tài liệu được quét với cấu hình quét: quét màu, độ phân giải 300 dpi;
 - + Lớp dưới là kết quả nhận dạng được ghép chính xác với lớp ảnh ở trên;
 - + Kích thước trung bình không quá 100 kb/trang A4.
 - Ứng dụng công nghệ có thể nhận dạng các cấu trúc logic, cách dàn trang cũng như các định dạng khác nhau trong tài liệu nhiều trang, ví dụ như: mục lục, đầu trang, chân trang, chú thích, chú thích bảng, chú thích ảnh, số trang... Khi kết quả nhận dạng được lưu dưới dạng Microsoft Word, các định dạng trên được nhìn nhận như những đối tượng tương ứng trong Word chứ không đơn thuần chỉ là các khối văn bản. Người dùng sẽ không mất, hoặc mất rất ít thời gian cho việc biên tập lại kết quả nhận dạng;
 - Kiến trúc hệ thống cho phép phân tải nhận dạng trên nhiều máy chủ để đáp ứng yêu cầu xử lý một số lượng lớn tài liệu trong thời gian ngắn;
 - Cho phép thêm các máy chủ xử lý nếu cần thiết tăng tốc độ nhận dạng để đáp ứng yêu cầu về tiến độ. Để giải quyết khối lượng dữ liệu lớn yêu cầu tốc độ bóc tách dữ liệu và xuất file pdf 2 lớp đạt 100 trang tài liệu trên 01 phút;
 - Cho phép cấu hình số tiến trình xử lý (processes); số nhân của chip (cores) sử dụng, lịch xử lý trên mỗi máy chủ xử lý để đảm bảo tận dụng tối đa năng lực, tài nguyên của các máy chủ.
 - Cho phép thiết lập cơ chế chạy tự động toàn bộ mục lục, toàn phong hoặc chọn từng hồ sơ.
 - Cho phép tự động nhận dạng và bóc tách thông tin từ các vùng thông tin quan trọng trong văn bản như: số ký hiệu, cơ quan ban hành, trích yếu văn bản, ngày ban hành;
 - Cho phép nhận dạng toàn văn văn bản, kết xuất ra nội dung text. Các nội dung này sẽ được lưu vào CSDL để phục vụ việc quản lý, tìm kiếm và truy xuất;
 - Có giao diện theo dõi hoạt động của hệ thống, trạng thái của từng yêu cầu nhận dạng;
 - Có giao diện để cấu hình các tiến trình xử lý (luồng công việc) với các thông tin sau:
 - + Nguồn ảnh đầu vào;
 - + Ngôn ngữ xử lý: hệ thống tự động nhận dạng toàn bộ ngôn ngữ trên văn bản;

+ Cách phân tách các tài liệu: Hệ thống có thể tách bộ tự động với độ chính xác trên 98% (Scan toàn bộ hồ sơ thành 01 file, hệ thống tự động cắt ra mỗi văn bản là 01 file);

- Có cơ chế cho phép thiết lập độ ưu tiên xử lý theo từng luồng công việc, có thể hủy yêu cầu nhận dạng nếu cần thiết;

- Có giao diện theo dõi hoạt động của hệ thống, trạng thái của từng yêu cầu nhận dạng;

3.3.5 Quy trình thực hiện số hóa tài liệu tại Lưu trữ lịch sử của Đảng



Diễn giải quy trình:

Số hoá tài liệu lưu trữ có thời hạn bảo quản vĩnh viễn. Việc số hóa được thực hiện theo các bước sau:

Bước 1. Chuẩn bị tài liệu

a) Lựa chọn tài liệu

Tài liệu lưu trữ cho số hoá để lập bản sao bảo hiểm bao gồm tất cả các tài liệu lưu trữ giấy có thời hạn bảo quản vĩnh viễn trong Lưu trữ lịch sử. Việc lựa chọn tài liệu thực hiện theo quy định của cơ quan, tổ chức. Lựa chọn tài liệu theo thứ tự ưu tiên gồm: Giá trị thông tin, tình trạng vật lý, tần suất khai thác tài liệu (trong đó giá trị thông tin là yếu tố quyết định, tình trạng vật lý và tần suất khai thác tài liệu là yếu tố bổ sung).

b) Giao nhận tài liệu giữa bộ phận lưu trữ và bộ phận số hoá tài liệu theo quy định của cơ quan, tổ chức.

Khi giao nhận phải:

- Kiểm đếm số lượng hồ sơ, số lượng tài liệu, số lượng tờ, số lượng trang trong từng hồ sơ;

- Lập Biên bản giao nhận hồ sơ, tài liệu kèm Danh mục thống kê số lượng tài liệu (gồm các thông tin: Hồ sơ, số đơn vị bảo quản, số tài liệu, số tờ, số trang).

c) Xử lý tài liệu trước khi số hoá

- Căn cứ tình trạng vật lý của tài liệu, có thể phải xử lý tài liệu trước khi số hoá như: Bóc, tách từng tờ tài liệu nếu tài liệu được đóng quyển; loại bỏ ghim, kẹp (nếu có);

- Làm phẳng khi tài liệu bị nhăn, gấp nếp hoặc phục chế nếu có khi tài liệu có tình trạng vật lý kém (cháy xém, rách nát, hư hỏng nặng, mốc, mờ chữ, mối mọt, ghim sắt gỉ lâu ngày, giấy ố vàng, tài liệu bị nhiễm axit nặng).

Bước 2. Thực hiện số hoá

a) Công tác chuẩn bị

- Vệ sinh máy quét bằng cồn và giẻ sạch.

- Kiểm tra nguồn điện, bật thiết bị lưu điện, bật máy quét, máy tính.

b) Số hoá tài liệu

Yêu cầu quét:

- Quét lần lượt từng trang tài liệu theo thứ tự trong hồ sơ (theo chiều đứng của chữ), không xáo trộn trật tự sắp xếp trong tài liệu;

- Đối với tài liệu có tình trạng vật lý tốt, đặt tài liệu trong máy quét cuộn (hay máy quét cuộn - Sheet feed scanner) để máy tự động quét.

- Đối với tài liệu có kích thước không đồng đều, mỏng, giòn dễ rách, đặt nằm từng trang tài liệu trong máy quét phẳng (hay máy quét tĩnh - Flatbed scanner) để quét;

- Đối với tài liệu có kích cỡ vượt quá khổ: Nếu không có máy quét tương ứng với khổ tài liệu có thể sử dụng máy quét phẳng khổ A4 (hoặc A3 nếu có) và chia tài liệu gốc thành từng phần A4 (hoặc A3 nếu có) và quét lần lượt từ trái qua phải, từ trên xuống dưới theo chiều kim đồng hồ, các phần liền kề gối lên nhau ít nhất 10 mm. Số thứ tự của ảnh được xác định theo số thứ tự quét.

Đặt chế độ ảnh trước khi thực hiện lệnh quét:

- Định dạng tệp tin (Format File): Quét đồng thời 2 định dạng Tag Image File - TIFF (.tif, phiên bản 6.0 trở lên) và Portable Document Format - PDF hoặc PDF/A (.pdf).

* TIFF: Phải là một trang, không chấp nhận tệp tin có nhiều trang.

* PDF, PDF/A: Tệp tin có thể có một trang hoặc nhiều trang tương ứng với số trang của tài liệu lưu trữ giấy.

+ Độ phân giải (Resolution): 300 dpi trở lên (hoặc 300 ppi), 600 dpi đối với tài liệu có khổ giấy nhỏ hơn A4 hoặc cỡ chữ văn bản dưới 8;

- Chế độ màu sắc (Colour):

* Nếu tài liệu giấy có các thành phần được tô màu hoặc trên nền giấy màu: Ảnh màu (Colour); độ sâu màu (Bit-depth) tối thiểu là 24 bit (8 bit cho mỗi kênh màu); không gian màu (Colour space) là sRGB.

* Nếu tài liệu giấy không có các thành phần được tô màu hoặc trên nền giấy không màu: Độ sâu bit (Bit depth) là 8 bit thang độ xám (greyscale); không gian màu (Colour space) là thang độ xám (greyscale).

+ Độ nén (Compression):

* Đối với định dạng TIFF: Chế độ không nén (Uncompressed, None, Type 0);

* Đối với định dạng PDF hoặc PDF/A: Có thể sử dụng hệ thống mã hoá lỗi JPEG 2000 phần 1 nén không mất dữ liệu (LZW) với lượng nén ít nhất với chất lượng hình ảnh tối đa.

+ Tỷ lệ quét (Scanning ratio): 100% (tức là kích thước gốc), tỷ lệ ảnh 1:1.

+ PDF, PDF/A có khả năng tìm kiếm (Searchability): Sử dụng nhận dạng ký tự quang học (OCR) để tìm kiếm tài liệu theo ngôn ngữ tiếng Việt (OCR tuân thủ PDF hoặc PDF/A).

Ngoài ra, tùy theo tình trạng vật lý của tài liệu để điều chỉnh độ tương phản sáng tối, độ bóng của chữ, mức độ màu, chất lượng hình ảnh... hoặc có thể đặt chế độ mặc định theo máy.

Lưu ý:

+ Việc quét tài liệu theo tệp tin số định dạng (.tif) là cho mục đích bảo hiểm và do cơ quan quản lý tài liệu lưu trữ quyết định.

+ Thiết bị quét có thể hỗ trợ 1 lần quét có thể xuất (lưu) đồng thời 2 tệp tin số định dạng Tag Image File - TIFF (.tif) và định dạng Portable Document Format - PDF hoặc PDF/A (.pdf).

+ Thiết bị quét có thể chưa hỗ trợ định dạng PDF/A, định dạng PDF sẽ được chấp nhận để số hoá với đầy đủ thông số kỹ thuật và sau đó sử dụng các phần mềm thương mại để chuyển đổi từ định dạng PDF sang định dạng PDF/A.

+ Trường hợp tài liệu lưu trữ đã được số hoá theo tệp tin số định dạng Portable Document Format (.pdf) và cập nhật dữ liệu thông qua việc thu thập (giao nộp) từ văn thư cơ quan, lưu trữ cơ quan hoặc trong quá trình hoạt động.

c) Lưu tên tệp tin số

Tên tệp tin được đặt thống nhất theo ký hiệu, mã định danh để bảo đảm là duy nhất, có khả năng tham chiếu và nhận diện mối liên hệ giữa tài liệu lưu trữ số với tài liệu lưu trữ giấy

- Đối với tệp tin số định dạng (.pdf), mỗi tệp tin tương ứng với một tài liệu giấy (tài liệu một trang hoặc nhiều trang). Có thể đặt tên tệp tin (.pdf) gồm các thông tin:

Mã định danh của cơ quan quản lý tài liệu; phong số (nhập đủ 3 số); mục lục số (nhập đủ 2 số); đơn vị bảo quản số (nhập đủ 4 số và ký tự số trùng nếu có); số thứ tự tài liệu trong đơn vị bảo quản (nhập đủ 3 số). Các thông tin nếu chưa đủ số thì thêm các số 0 ở phía trước và được nối với nhau bởi dấu gạch ngang không cách (-).

Ví dụ: A38-011-07-0123a-001.pdf

Ngoài ra, tên tệp tin có thể bổ sung các thông tin siêu dữ liệu, như: Ký hiệu tên loại văn bản; số văn bản; năm văn bản...

Ví dụ: A38-011-07-0123-001-BC-0001-1998.pdf

- Đối với tệp tin số định dạng (.tif), mỗi tệp tin tương ứng với một trang tài liệu giấy. Có thể đặt tên tệp tin (.tif) gồm các thông tin: Mã định danh của cơ quan quản lý tài liệu; phong số (nhập đủ 3 số); mục lục số (nhập đủ 2 số); đơn vị bảo quản số (nhập đủ 4 số và ký tự số trùng nếu có); trang số (nhập đủ 3 số). Các thông tin nếu chưa đủ số thì thêm các số 0 ở phía trước và được nối với nhau bởi dấu gạch ngang không cách (-).

Ví dụ: A38-011-07-0123-001.tif

A38-011-07-0123-002.tif

A38-011-07-0123-003.tif

..."

d) Lưu tệp tin số vào thư mục

- Sau khi số hoá xong toàn bộ các trang tài liệu của một hồ sơ và lưu tên tệp tin số (file), thực hiện việc xuất các tệp tin số (file) vào thư mục hồ sơ;

- Thư mục hồ sơ: Đặt tên thư mục hồ sơ theo thông tin của địa chỉ bảo quản;

Cách thức tạo lập thư mục hồ sơ theo địa chỉ bảo quản: Tên thư mục hồ sơ được đặt theo quy tắc gồm các thông tin: Mã định danh của cơ quan quản lý tài liệu, Phong số, Mục lục số, Đơn vị bảo quản số. Các thông tin được nối với nhau bởi dấu gạch ngang không cách (-), cụ thể:

+ Phong số: Nhập đủ 3 số, nếu phong số chưa đủ 3 số thì thêm các số 0 ở phía trước;

+ Mục lục số: Nhập đủ 2 số, nếu mục lục số chưa đủ 2 số thì thêm các số 0 ở phía trước;

+ Đơn vị bảo quản số: Nhập đủ 4 số, nếu đơn vị bảo quản số chưa đủ 4 số thì thêm các số 0 ở phía trước

Ví dụ: A38-011-07-0123

- Trong thư mục hồ sơ, có thể tạo lập các thư mục tài liệu theo danh mục tài liệu bên trong hồ sơ và xuất các tệp tin số (file) theo thư mục tài liệu;

Cách thức tạo lập thư mục tài liệu trong thư mục hồ sơ: Tên thư mục tài liệu được đặt theo quy tắc gồm các thông tin: Mã định danh của cơ quan quản lý tài liệu, Phong số, Mục lục số, Đơn vị bảo quản số, Số thứ tự tài liệu trong đơn vị bảo quản. Các thông tin được cách nhau bởi dấu gạch ngang không cách (-). Cụ thể:

- Phong số: Nhập đủ 3 số, nếu chưa đủ 3 số thì thêm các số 0 ở phía trước;

- Mục lục số: Nhập đủ 2 số, nếu chưa đủ 2 số thì thêm các số 0 ở phía trước;
- Đơn vị bảo quản số: Nhập đủ 4 số, nếu chưa đủ 4 số thì thêm các số 0 ở phía trước.

- Số thứ tự tài liệu được đơn vị bảo quản: Nhập đủ 3 số, nếu chưa đủ 3 số thì thêm các số 0 ở phía trước.

Ví dụ: Mã định danh của cơ quan quản lý A38, Hồ sơ có số đơn vị bảo quản 123, thuộc mục lục số 7, phong số 11, có 4 văn bản, 12 trang. Trong đó văn bản thứ nhất có 3 trang, văn bản thứ 2 có 4 trang, văn bản thứ 3 có 3 trang, văn bản thứ 4 có 2 trang. Đặt tên các thư mục tài liệu như sau:

- Tài liệu 1 có 3 trang, đặt tên thư mục như sau: A38-011-07-0123-001

+ Thư mục tài liệu 1 lưu các tệp (file) tương ứng với trang của tài liệu:

A38-011-07-0123-001.tif

A38-011-07-0123-002.tif

A38-011-07-0123-003.tif

- Tài liệu 2 có 4 trang, đặt tên thư mục như sau: A38-011-07-0123-002

+ Thư mục tài liệu 2 lưu các tệp (file) tương ứng với trang của tài liệu:

A38-011-07-0123-004.tif

A38-011-07-0123-005.tif

A38-011-07-0123-006.tif

A38-011-07-0123-007.tif

- Tài liệu 3 có 3 trang, đặt tên thư mục như sau: A38-011-07-0123-003

+ Thư mục tài liệu 3 lưu các tệp (file) tương ứng với trang của tài liệu:

A38-011-07-0123-008.tif

A38-011-07-0123-009.tif

A38-011-07-0123-010.tif

- Tài liệu 4 có 2 trang, đặt tên thư mục như sau: A38-011-07-0123-004

+ Thư mục tài liệu 4 lưu các tệp (file) tương ứng với trang của tài liệu:

A38-011-07-0123-011.tif

A38-011-07-0123-012.tif

Các thư mục tài liệu được đặt trong thư mục hồ sơ A38-011-07-0123.

Lưu ý: Khuyến khích lưu tệp tin số định dạng PDF (.pdf) cùng thư mục với tệp tin số định dạng (tif)"

Bước 3. Kiểm tra chất lượng số hoá, bàn giao tài liệu

a) Kiểm tra chất lượng số hoá

- Phải kiểm tra chất lượng màn hình trước khi kiểm tra hình ảnh của tài liệu, đặt chế độ phân giải màn hình phù hợp, chế độ xem ảnh ở tỉ lệ 1:1 (100%).

- Tiến hành kiểm tra hình ảnh trên màn hình và đối chiếu với tài liệu gốc theo một số thuộc tính sau:

+ Kiểm tra tính đầy đủ: Kiểm tra các ảnh của tài liệu theo thứ tự sắp xếp trong tài liệu. Chú ý kiểm tra các tờ tài liệu có hai mặt, các trang tài liệu quá khổ...

+ Kiểm tra hình ảnh: Hình ảnh tài liệu bị xiên, lệch; chữ bị cắt ở các lề; độ tương phản, chữ bị mất nét; sự khác biệt so với tài liệu gốc...

+ Kiểm tra các thông số kỹ thuật theo chế độ quét đã thiết lập.

b) Quét lại các tài liệu chưa đạt yêu cầu (nếu có)

Trang tài liệu quét lại phải được kiểm tra, đặt tên tệp tin số (file) theo quy định nêu trên.

c) Bàn giao tài liệu giấy

Việc bàn giao tài liệu giữa bộ phận số hoá và bộ phận lưu trữ theo quy định của cơ quan, đơn vị.

Khi bàn giao phải:

- Sắp xếp tài liệu theo đúng trật tự sắp xếp tài liệu trong hồ sơ.

- Căn cứ Biên bản giao nhận và Danh mục thống kê số lượng tài liệu để kiểm đếm từng trang tài liệu trong hồ sơ, giao nộp số lượng hồ sơ.

- Lập biên bản giao tài liệu (gồm các thông tin: số đơn vị bảo quản, số tài liệu, số tờ, số trang, thứ tự sắp xếp của tài liệu trong hồ sơ).

Bước 4. Tập trung dữ liệu

- Sau mỗi ngày số hoá, phải sao lưu toàn bộ tệp tin số (gồm tệp tin số *Tập tin chủ (Master File) theo tiêu chuẩn định dạng TIFF (.tif)* và thư mục) từ máy tính dùng để số hoá sang lưu trữ trong phương tiện lưu trữ điện tử để lưu trữ và bảo quản lâu dài. Tiến hành kiểm tra thông tin dữ liệu đã sao lưu trên phương tiện lưu trữ điện tử và xoá dữ liệu ảnh lưu trong máy tính dùng để số hoá;

- Căn cứ Danh mục thống kê số lượng tài liệu để kiểm đếm số lượng thư mục theo hồ sơ, số lượng thư mục tài liệu trong hồ sơ, số lượng tệp tin số trong hồ sơ; kiểm tra chất lượng hình ảnh và cách đặt tên;

- Lập danh mục thống kê số lượng tệp tin số theo hồ sơ (gồm các thông tin: Số đơn vị bảo quản, tổng số trang, tổng số lượng tệp tin số, tổng dung lượng theo hồ sơ, địa chỉ thư mục theo hồ sơ, thời gian số hoá);

Bước 5. Cập nhật dữ liệu số hoá vào cơ sở dữ liệu

a) Xuất tệp tin số định dạng TIF sang tệp tin số định dạng PDF

Từ tệp tin số định dạng (.tif) có thể tạo lập bản sao tài liệu thành tệp tin số định dạng (.pdf) bằng cách chuyển đổi định dạng (hay di trú) mà không phải số hoá tài liệu.

+ Cách thức chuyển đổi định dạng thực hiện như sau: Sử dụng phần mềm thương mại để chuyển đổi tệp tin số định dạng (.tif) sang tệp tin số định dạng (.pdf). Trường hợp tài liệu có 1 trang thì chuyển đổi 1 tệp tin số định dạng (.tif) tương ứng thành 1 tệp tin số định dạng (.pdf). Trường hợp tài liệu có nhiều trang thì chuyển đổi mỗi tệp tin số định dạng (.tif) tương ứng thành một tệp tin số định dạng (.pdf) và ghép (gộp) các tệp tin số định dạng (.pdf) thành một tệp tin số định dạng

(.pdf) tương ứng với tài liệu giấy có nhiều trang.

+ Đặt tên tệp tin số: Tên tệp tin số (file) định dạng (.pdf) gồm các thông tin: Mã định danh của cơ quan quản lý tài liệu; phong số (nhập đủ 3 số); mục lục số (nhập đủ 2 số); đơn vị bảo quản số (nhập đủ 4 số và ký tự số trùng nếu có); số thứ tự tài liệu trong đơn vị bảo quản (nhập đủ 3 số). Các thông tin nếu chưa đủ số thì thêm các số 0 ở phía trước và được nối với nhau bởi dấu gạch ngang không cách (-).

Ví dụ: A38-011-07-0123-001.pdf

Ngoài ra, tên tệp có thể bổ sung các thông tin siêu dữ liệu mô tả tài liệu lưu trữ giấy để tra cứu, như: Ký hiệu tên loại văn bản; số văn bản; năm văn bản...

Ví dụ: A38-011-07-0123-001-BC-0001-1998.pdf

Lưu ý:

+ Sử dụng ứng dụng phần mềm nhận dạng ký tự quang học (OCR) để hỗ trợ tìm kiếm tài liệu (văn bản) theo ngôn ngữ tiếng Việt (OCR tuân thủ PDF, PDF/A).

+ Có thể sử dụng các ứng dụng phần mềm thương mại để chuyển đổi định dạng PDF sang định dạng PDF/A".

b) Sử dụng chữ ký số để chứng thực lưu trữ đối với tài liệu lưu trữ

Sử dụng chữ ký số của lưu trữ để chứng thực lưu trữ đối với tệp tin số (file) định dạng PDF (.pdf).

c) Cập nhật cơ sở dữ liệu

- Cập nhật thông tin cấp 2 của tài liệu vào cơ sở dữ liệu quản lý tài liệu lưu trữ;
- Tạo siêu dữ liệu liên kết hay gắn tệp tin số (file) định dạng PDF (.pdf) đã tích hợp chữ ký số vào cơ sở dữ liệu quản lý tài liệu lưu trữ;
- Kiểm tra lại các thông tin đã cập nhật và thông tin tệp toàn văn tài liệu.

Lưu ý:

+ Có thể sử dụng phần mềm chuyên dụng để chuyển đổi tệp tin số (file) định dạng PDF (.pdf) thành tệp tin số (file) định dạng PDF/A (.pdf) để hỗ trợ tra cứu;

+ Có thể sử dụng các phần mềm chuyên dụng để tự động cập nhật thông tin cấp 2 vào cơ sở dữ liệu quản lý tài liệu lưu trữ.

Bước 6. Lập hồ sơ về việc số hoá phonh/khôi tài liệu

Hồ sơ gồm các văn bản:

- Kế hoạch về số hoá phonh/khôi tài liệu;
- Biên bản giao nhận (kèm theo Danh mục thống kê số lượng tài liệu, Danh mục thống kê số lượng tệp tin số theo hồ sơ);
- Các loại văn bản, giấy tờ khác có liên quan.

3.3.6 Xuất tệp tin số phục vụ khai thác, sử dụng

Căn cứ vào mục đích khai thác, sử dụng mà áp dụng các phương pháp xuất tệp tin số theo các định dạng tương thích để tăng hoặc giảm dung lượng tệp tin số. Giảm dung lượng tệp tin số để tăng tốc độ truy cập tệp tin số phù hợp với hạ tầng công nghệ thông tin hiện tại. Sử dụng các công cụ, phần mềm dịch vụ hỗ trợ để tăng

hoặc giảm dung lượng tệp tin số.

Các phương pháp xuất tệp tin số (*Tập tin chủ định dạng TIFF (.tif)* hoặc *Tập tin giao dịch định dạng PDF (.pdf)*) để phục vụ cho mục đích khai thác, sử dụng như sau:

- Để đọc, in ấn, tra tìm trong cơ sở dữ liệu trên mạng máy tính của cơ quan, tổ chức:

- + Thay đổi độ phân giải phù hợp cho bản in (150 dpi);
- + Độ nén (JPEG lossless);
- + Khổ giấy in chính xác (tỉ lệ 1:1);
- + Đầy đủ thông chữ mà các ký tự trong tài liệu sử dụng.

- Để đọc trên màn hình máy tính (không để in hoặc in rất mờ) trong cơ sở dữ liệu trên mạng thông tin diện rộng của Đảng, đăng lên từng tin điện tử website...:

- + Giảm độ phân giải (72dpi);
 - + Tăng độ nén ảnh;
 - + Giảm khổ giấy (đổi A4 thành A5, A6 hoặc theo tỉ lệ 1:2, 1:4);
 - + Chuyển đổi ảnh màu thành đen trắng..,
 - + Sử dụng chữ ký số để chứng thực lưu trữ.
- Để chế bản khổ lớn phục vụ triển lãm, công bố,...:
- + Thay đổi độ phân giải phù hợp cho bản in (450-600 dpi);
 - + Khổ giấy in (tỉ lệ 2:1,3:1);
 - + Sử dụng chữ ký số để chứng thực lưu trữ.

3.4. Yêu cầu Ổ cứng lưu trữ tài liệu số hóa

- Nhà thầu cam kết cung cấp ổ cứng lưu trữ tài liệu số hóa đảm bảo chưa qua sử dụng, mới 100%, năm sản xuất 2025 trở lại đây, dung lượng lưu trữ > 08TB, có CO/CQ hoặc tài liệu tương đương chứng minh nguồn gốc, xuất xứ.

3.5. Yêu cầu về bảo hành

- Thời gian bảo hành: 12 tháng kể từ ngày bàn giao, nghiệm thu.

- Nội dung bảo hành:

+ Thực hiện cập nhật, chỉnh sửa các sai sót, lỗi phát sinh trong quá trình sử dụng tài liệu, dữ liệu số hóa.

- Cơ chế giải quyết các sai sót, lỗi phát sinh (nếu có) trong quá trình sử dụng dữ liệu trong thời gian bảo hành:

+ Hỗ trợ trực tiếp tại các địa điểm của Chủ đầu tư.

Mọi chi phí khắc phục các nội dung bảo hành trong thời gian bảo hành do nhà thầu chịu.

4. Giải pháp và phương pháp luận:

Nhà thầu chuẩn bị đề xuất giải pháp, phương pháp luận tổng quát thực hiện dịch vụ theo các nội dung quy định tại Chương này, gồm các phần như sau:

1. Giải pháp và phương pháp luận;

2. Kế hoạch công tác.

5. Quy định về kiểm tra, nghiệm thu sản phẩm:

Theo quy định pháp luật hiện hành.