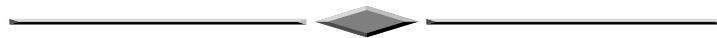


ГЕОДЕЗИЯ И МАРКШЕЙДЕРИЯ



УДК 528.48

DOI 10.33764/2411-1759-2025-30-4-5-14

Автоматизация процессов BIM-моделирования по данным проектной документации

Д. М. Искаков¹

¹ Сибирский государственный университет геосистем и технологий, г. Новосибирск, Российская Федерация

e-mail: d.iskakov@svarog.ru

Аннотация. В статье рассматривается процесс автоматизации обработки данных проектной документации и их интеграции в BIM-модели. Основное внимание уделено этапу подготовки данных, который традиционно считается наиболее трудоемким и ресурсоемким. Цель исследования – автоматизация процессов BIM-моделирования по данным проектной документации. Основными задачами являются автоматизация извлечения данных из PDF-чертежей, их последующая обработка и преобразование в формат, пригодный для использования в системах BIM. Результаты показали, что разработанный инструмент способен существенно сократить время обработки данных и повысить точность их интеграции. Основные выводы, сделанные по результатам работы, заключаются в том, что предложенный инструмент значительно упрощает и ускоряет такие процессы, как преобразование данных из технической документации, извлечение текстовой и табличной информации, а также проверка их на полноту и соответствие. Это, в свою очередь, повышает эффективность создания каталогов и миникаталогов компонентов трубопроводных систем, что способствует повышению точности и сокращению времени на проектирование. Это особенно важно в условиях обработки больших объемов данных, где традиционные методы обработки могут оказаться трудоемкими и менее точными.

Ключевые слова: Python, оптическое распознавание текста, автоматизация процессов, PDF-чертежи, информационное моделирование зданий, извлечение данных, интеграция данных

Для цитирования:

Искаков Д. В. Автоматизация процессов BIM-моделирования по данным проектной документации // Вестник СГУГиТ. – 2025. – Т. 30, № 4. – С. 5–14. – DOI 10.33764/2411-1759-2025-30-4-5-14

Введение

Современные технологии информационного моделирования зданий (BIM) становятся ключевым инструментом в управлении жизненным циклом объектов, начиная от проектирования и заканчивая эксплуатацией. Интеграция данных проектной документации

в BIM-модели часто связана с рядом сложностей, включая большие объемы данных, их структурирование и обработку. В этом контексте использование языка программирования для автоматизации процессов обработки данных и их интеграции в BIM-модели представляет собой перспективное направление исследований.

Цель данного исследования заключается в автоматизации процессов BIM-моделирования по данным проектной документации. Основными задачами являются автоматизация извлечения данных из PDF-чертежей, их последующая обработка и преобразование в формат, пригодный для использования в системах BIM, таких как AutoCAD Plant 3D. Основное внимание уделяется автоматизации извлечения информации из таблиц, содержащих данные о компонентах трубопроводных систем, представленных в изометрических чертежах.

Теоретическая значимость исследования заключается в развитии методов автоматизации обработки проектных данных для использования в BIM моделировании, что позволит существенно сократить время на обработку данных и снизить количество ошибок при интеграции. Практическая значимость заключается в разработке программного инструмента, который может быть использован в реальных строительных проектах для повышения эффективности процессов моделирования и управления строительными объектами.

Несмотря на значительное количество исследований в области применения лазерного сканирования и BIM, остается ряд нерешенных проблем. Например, автоматизация процесса обработки больших объемов данных и их интеграции в BIM-модели все еще требует разработки эффективных алгоритмов и инструментов [1–3]. Кроме того, проблема точности распознавания и обработки данных, полученных из PDF-чертежей, особенно в контексте сложных инженерных объектов, также требует дальнейшего изучения [4–6].

Среди наиболее известных исследований в данной области можно выделить работы, посвященные автоматизированному созданию BIM-моделей на основе облаков точек, полученных с помощью НЛС [7–9]. Эти исследования демонстрируют значительные достижения в области моделирования сложных объектов, однако не охватывают все аспекты интеграции данных и автоматизации процессов, что подчеркивает актуальность и необходимость дальнейших разработок.

Доступные публикации отечественных исследователей, затрагивающие BIM-технологии, преимущественно ориентированы на

общие вопросы их внедрения и применения в строительной отрасли [10, 11]. Следует отметить, что специализированные исследования, посвященные автоматизации обработки данных НЛС и их интеграции в BIM-модели, в российских научных источниках практически отсутствуют.

Таким образом, представляемая работа нацелена на решение указанных проблем путем разработки эффективного инструмента, который позволит оптимизировать процессы обработки данных изометрических чертежей и интеграции их в BIM-модели, что, в свою очередь, повысит эффективность проектирования и управления строительными проектами.

Анализ последних исследований показывает значительный интерес к вопросам интеграции данных НЛС в BIM. В частности, в работах [2, 3] рассматриваются подходы к автоматизированному созданию BIM-моделей на основе облаков точек, а исследования [4, 8] подчеркивают важность оптимизации этих процессов для повышения точности и снижения затрат. Тем не менее, многие аспекты, связанные с автоматизацией обработки данных и их интеграцией в BIM, остаются нерешенными, что делает данное исследование актуальным и востребованным.

Методы и материалы

Процесс создания BIM-модели состоит из нескольких важных этапов [12], каждый из которых обеспечивает поэтапное и логичное преобразование данных, представленных в чертежах и спецификациях, в полноценную цифровую модель. Методика построения модели представлена на рис. 1.



Рис. 1. Основные этапы построения модели

Особое внимание в данном исследовании уделяется автоматизации этапа «Подготовка данных к обработке», так как этот этап является наиболее трудоемким.

В рамках проведенного исследования был разработан программный инструмент на

языке Python 3.12, который автоматизирует этап «Подготовка данных к обработке». В исследовании применялись изометрические чертежи нефтепровода, фрагмент изометрического чертежа в формате «.pdf» представлен на рис. 2.

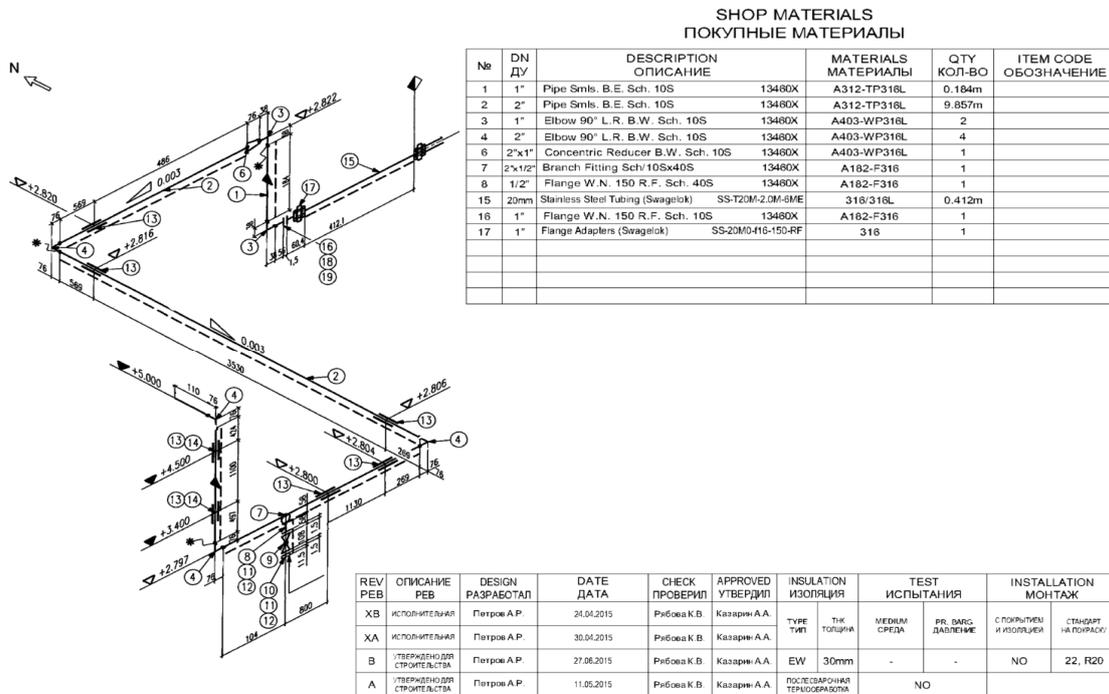


Рис. 2. Фрагмент изометрического чертежа в формате «.pdf»

Для работы с текстовой информацией используются два подхода: если данные представлены в машиночитаемом формате, применяются инструменты «pdfplumber», «PyMuPDF (fitz)» и «PDFMiner», которые позволяют анализировать структуру документа и извлекать текстовые блоки с сохранением их логической организации. В случае, когда текст представлен в виде изображений, используется «Tesseract OCR», «EasyOCR» и «PaddleOCR», предварительно оптимизированные методами бинаризации и фильтрации шумов для повышения точности распознавания.

Основные методы и материалы, использованные в работе, включают несколько ключевых модулей, каждый из которых выполняет поставленные оператором задачи, обеспечивая целостность и последовательность обработки данных. Структура ключевых модулей программы представлена на рис. 3.

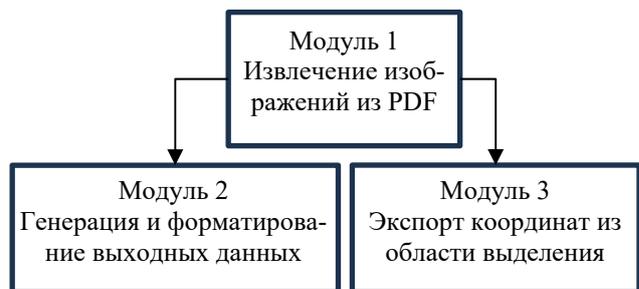


Рис. 3. Структура ключевых модулей программы

Модуль 1 является основным компонентом программы, который отвечает за реализацию последовательного извлечения изображений из PDF-документов. Этот процесс осуществляется с помощью библиотеки «Pdfplumber», которая обеспечивает точное извлечение визуальных данных из чертежей [13]. Принцип работы библиотеки «Pdfplumber» представлен на рис 4.

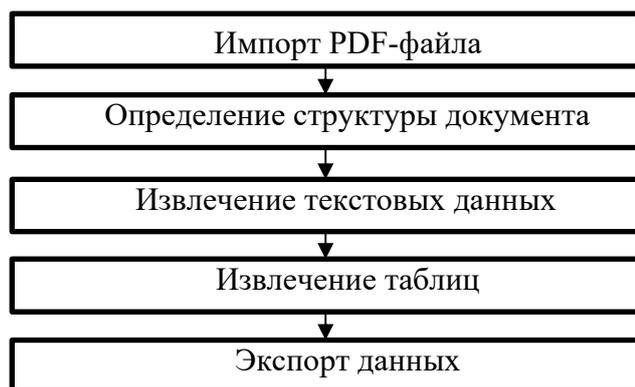


Рис. 4. Принцип работы библиотеки «Pdfplumber»

Извлеченные изображения затем подвергаются дальнейшей обработке, включающей фильтрацию шума, повышение контрастности и преобразование в черно-белый формат. Эти этапы обработки реализуются с использованием библиотеки «Pillow», что позволяет существенно улучшить качество изображений перед применением OCR (оптического распознавания текста) [14].

Для распознавания текста на обработанных изображениях используется библиотека «Pytesseract», которая интегрирует движок Tesseract OCR. Этот метод позволяет эффективно извлекать текстовую информацию с изображений, особенно после их предварительной обработки [15]. Принцип работы библиотеки «Pytesseract» представлен на рис 5.

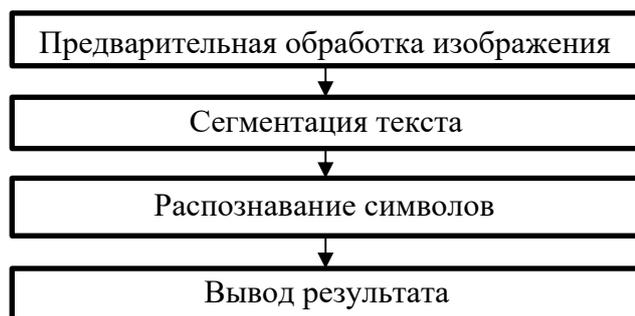


Рис. 5. Принцип работы библиотеки «Pytesseract»

Настройки программы, такие как пути к файлам и параметры обработки изображений, управляются с помощью библиотеки «Configparser», что обеспечивает гибкость и адаптив-

ность программы без необходимости изменения исходного кода. Модуль «Configparser», являющийся встроенным инструментом языка программирования Python, предназначен для эффективной работы с конфигурационными файлами, представленными в формате «.INI». Основным принципом работы модуля основан на организации данных в виде секций, содержащих пары «ключ – значение». Такой подход позволяет структурированно хранить параметры, обеспечивая логичное разделение настроек по функциональным блокам [16].

Модуль 2 отвечает за генерацию и форматирование выходных данных в формате «.xlsx». Данные, полученные в результате распознавания текста, структурируются и организуются в виде таблиц, которые затем сохраняются в Excel-файл с использованием библиотеки «Orepanxl» [17]. Этот процесс включает настройку ширины столбцов в соответствии с содержимым, что улучшает визуальное восприятие данных, а также удаление лишних пустых строк для обеспечения целостной структуры итогового документа.

Модуль 3 предоставляет графический интерфейс, позволяющий пользователю интерактивно выделять области на изображениях, из которых требуется извлечь текст. Этот интерфейс разработан на основе библиотеки «Tkinter» и позволяет пользователю точно определять и называть области изображения, которые затем сохраняются в виде координат [18]. Инструмент значительно упрощает дальнейшее использование данных, обеспечивая возможность автоматизации процесса выделения областей на изображениях. Это особенно важно при многократном использовании программы, когда необходимо работать с повторяющимися задачами.

Результаты

В результате проведенного исследования был разработан программный инструмент на языке Python 3.12, который успешно автоматизирует процесс обработки больших объемов данных проектной документации. Блок-схема процесса извлечения данных из PDF и их экспорта в Excel представлена на рис. 6.

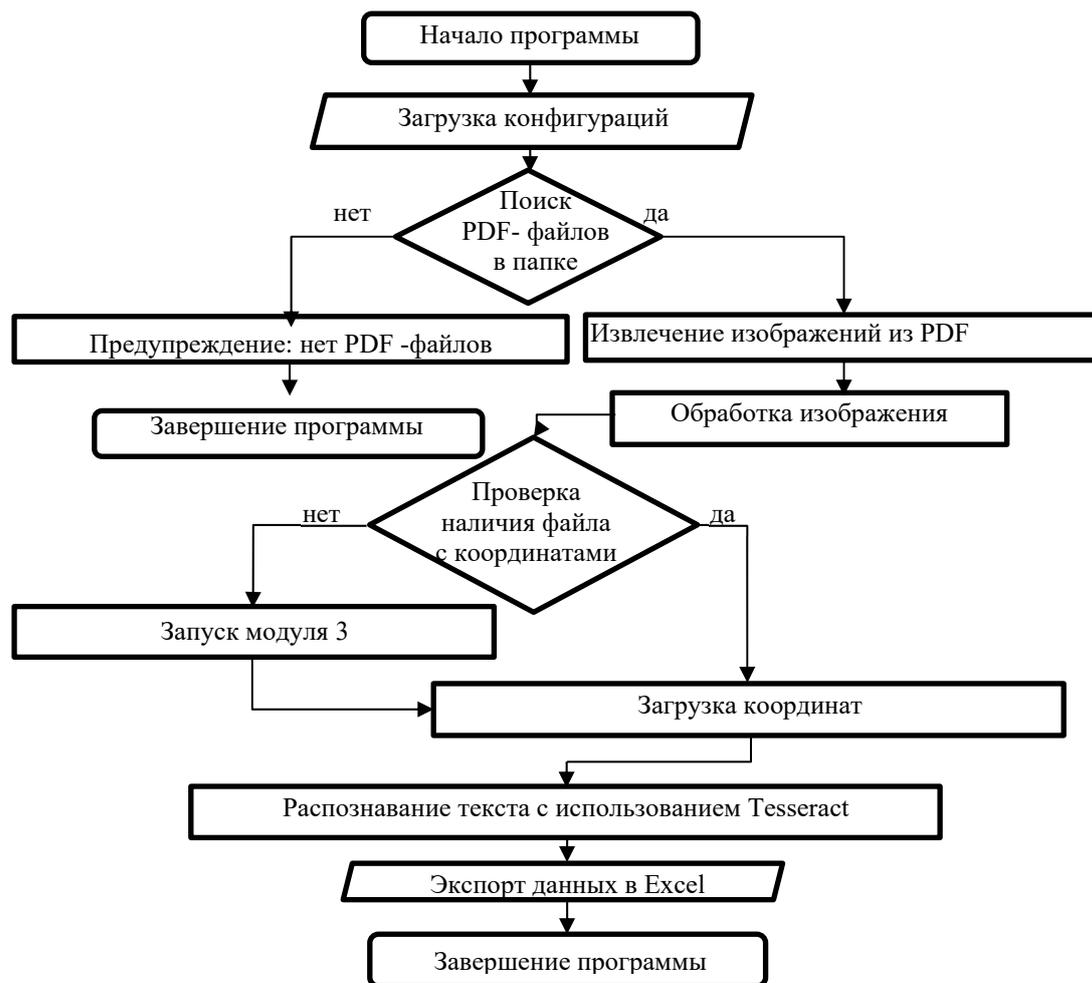


Рис. 6. Блок-схема процесса извлечения данных из PDF и их экспорта в Excel

Работа модуля «Экспорт координат из области выделения» представлена на рис. 7. Данный модуль предоставляет пользователю возможность интерактивного взаимодействия с изображениями, позволяя выделять необходимые области с высокой точностью.

SHOP MATERIALS ПОКУПНЫЕ МАТЕРИАЛЫ					
№	DN ДУ	DESCRIPTION ОПИСАНИЕ	MATERIALS МАТЕРИАЛЫ	QTY КОЛ-ВО	ITEM CODE ОБОЗНАЧЕНИЕ
1	1"	Pipe Smpls. B.E. Sch. 10S	13480X	A312-TP316L	0.184m
2	2"	Pipe Smpls. B.E. Sch. 10S	13480X	A312-TP316L	9.857m
3	1"	Elbow 90° L.R. B.W. Sch. 10S	13480X	A403-WP316L	2
4	2"	Elbow 90° L.R. B.W. Sch. 10S	13480X	A403-WP316L	4
6	2"x1"	Concentric Reducer B.W. Sch. 10S	13480X	A403-WP316L	1
7	1/2"	Branch Fitting Sch. 10S 40S	13480X	A182-F316	1
8	1/2"	Flange W.N. 150 R.F. Sch. 40S	13480X	A182-F316	1
15	20mm	Stainless Steel Tubing (Swagelok)	SS-20M2.0M-2ME	318/316L	0.412m
16	1"	Flange W.N. 150 R.F. Sch. 10S	13480X	A182-F316	1
17	1"	Flange Adapters (Swagelok)	SS-20M0416-150-4F	318	1

Рис. 7. Работа модуля «Экспорт координат из области выделения»

Пример координат для первой таблицы, получаемых с помощью модуля «Экспорт координат из области выделения», представлен на рис. 8.

ординат из области выделения», представлен на рис. 8.

Z1: (7933.0, 252.0), (9161.0, 490.0)
 R1-1: (6873.0, 536.0), (7039.0, 1743.0)
 R1-2: (7058.0, 532.0), (7224.0, 1812.0)
 R1-3: (7231.0, 529.0), (8618.0, 1858.0)
 R1-4: (8707.0, 545.0), (9218.0, 1728.0)
 R1-5: (9291.0, 541.0), (9580.0, 1822.0)
 R1-6: (9623.0, 541.0), (10216.0, 1717.0)

Рис. 8. Пример получаемых координат модулем «Экспорт координат из области выделения»

Программный инструмент тестировался на 20 файлах в формате «.pdf». Тесты показали следующие результаты:

- скорость обработки одного файла составила одну минуту;
- скорость обработки двадцати файлов составила десять минут;

– скорость выделения необходимых областей для данного примера составила шесть минут;
 – скорость переноса данных одного файла вручную занимает от двух до четырех минут.

Точность переноса текста в «.xlsx» формат автоматическим способом составила 80 %, что требует последующего ручного корректирования. Пример итогового файла Excel представлен на рис. 9.

№	DN	ОПИСАНИЕ	МАТЕРИАЛЫ	QTY	ИТЕМ CODE
4	E	HY	МАТЕРИАЛЫ	KON-BO	ОБОЗНАЧЕНИЕ
5	20	4m	Pipe Smls. B.E. Sch. 105 13460X	A312-TP316L	0.184m
6	o8	ra	Pipe Smls. B.E. Sch. 105 13460X	A312-TP316L	9.857m
7	4	a	Elbow 90° L.R. B.W. Sch. 105 13460X	A403-WP316L	2
8	6j	Га	Elbow 90° L.R. B.W. Sch. 105 13460X	A403-WP316L	4
9	7	412"	Concentric Reducer B.W. Sch. 105 13460X	A403-WP316L	—
10	8	20mm	Branch Fitting Sch/108x405 13460X	A182-F316	—
11	015°	a	Flange V.W.N. 150 R.F. Sch. 405 13460X	A182-F316	—
12	16°	K	Stainless Steel Tubing (Swagelok) SS-T20M-2.0M-6ME	316/316L	0.412m
13	a7		Flange W.N. 150 R.F. Sch. 105 13460X	A182-F316	—
14			Flange Adapters (Swagelok) SS-20M00-F16-150-RF	316	—

Рис. 9. Пример итогового файла Excel

На следующем этапе для построения модели требуется систематизировать текстовую информацию, а также провести тщательную проверку и исправление грамматических ошибок. Исправленный документ станет основой для быстрой и эффективной интеграции всех данных в миникаталог компонентов трубопровода программного комплекса AutoCAD Plant 3D. Каталоги используются для создания спецификаций, которые в дальнейшем применя-

ются для моделирования трубопроводных систем. Миникаatalogи работают аналогично основным каталогам, но могут быть более узко специализированными и содержать только нужные компоненты, что упрощает работу с часто используемыми элементами. Пример окна редактора миникаatalogа представлен на рис. 10. Пример подготовленной детали по полученным данным из перенесенного текста представлен на рис. 11.

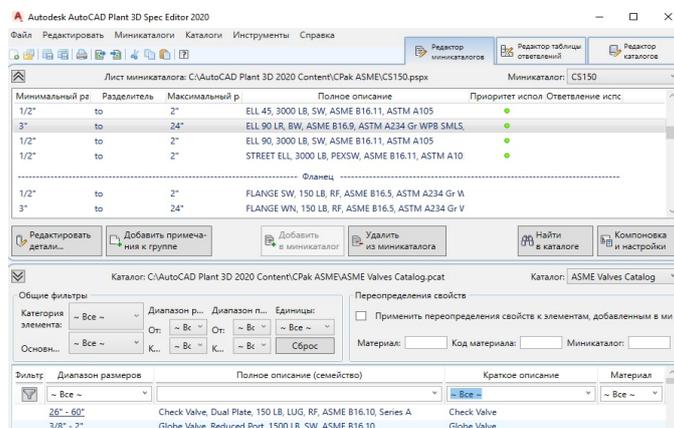


Рис. 10. Окно редактора миникаatalogа

Размер	Полное описание (семейство)	Совместимый	Изготовитель	Материал	Код	Полное описание (размер)
1/2"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
3/4"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
1"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
1 1/4"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
1 1/2"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
2"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
2 1/2"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
3"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
3 1/2"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
4"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
5"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
6"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
8"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
10"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
12"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
14"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
16"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...
18"	Elbow 90° L.R. B.W. Sch. 105 13460X	ASME B16.9		A403-WP316L		Elbow 90° L.R. B.W. Sch. 105 13...

Рис. 11. Пример подготовленной детали в AutoCAD Plant 3D

Созданный миникаталог предоставляет возможность оперативно моделировать определенный участок трубопровода в соответствии с изометрическим чертежом и всеми его спецификациями. Такой подход значительно упрощает процесс проектирования, обеспечивая точность и соответствие исходным данным. Пример модели, построенной по полученному миникаталогу, представлен на рис. 12.



Рис. 12. Пример модели, построенной по полученному миникаталогу в AutoCAD Plant 3D

Заключение

В ходе исследования была проведена работа по автоматизации процессов BIM моделирования по данным проектной документации. Для автоматизации этапа «Подготовка данных к обработке» был разработан про-

граммный инструмент, который успешно автоматизирует процесс обработки больших объемов данных проектной документации для дальнейшего использования в BIM.

Результаты исследования показывают, что внедрение автоматизированных инструментов может значительно улучшить процессы обработки данных. Это, в свою очередь, способствует более эффективной интеграции данных в BIM-моделирование и повышению точности проектных решений.

Точность распознавания текста в данном исследовании составила 80 %, что требует дальнейших улучшений. На данный момент все еще необходима коррекция полученных данных в интерактивном режиме. Это можно рассматривать как направление для дальнейших исследований, направленное на совершенствование алгоритмов распознавания текста и улучшение качества обработки изображений.

Созданный миникаталог на основе полученного текста предоставляет возможность оперативно моделировать определенный участок трубопровода в соответствии с изометрическим чертежом и всеми его спецификациями.

В качестве дальнейших исследований можно рассмотреть вариант работы с разными форматами данных, а также расширить возможности программного комплекса, что повысит эффективность и гибкость в решении практических задач.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Zaczek-Peplinska J., Pasik M., Adamek A., Kołakowska M., lapinski S. Monitoring technical conditions of engineering structures using the terrestrial laser scanning technology // Reports on Geodesy and Geoinformatics. – 2013. – № 95. – С. 1–10. – DOI 10.2478/rgg-2013-0008.
2. Wang J., Kutterer H., Fang X. On the detection of systematic errors in terrestrial laser scanning data // Journal of Applied Geodesy – 2012. – № 6. – С. 187–192. – DOI 10.1515/jag-2012-0025.
3. Błaszczyk M., Laska M., Sivertsen A., Jawak S. Combined use of aerial photogrammetry and terrestrial laser scanning for detecting geomorphological changes in Hornsund, Svalbard // Remote Sensing. – 2022. – № 14. – С. 123–131. – DOI 10.3390/rs14030601. – EDN COHIBP.
4. Liu J., Fu L., Cheng G., Li D. Automated BIM reconstruction of full-scale complex tubular engineering structures using terrestrial laser scanning // Remote Sensing. – 2022. – № 14. – С. 542–553. – DOI 10.3390/rs14071659. – EDN PDPXWC.
5. Dupuis J., Holst C., Kuhlmann H. Laser scanning based growth analysis of plants as a new challenge for deformation monitoring // Journal of Applied Geodesy. – 2016. – № 10. – С. 37–44. – DOI 10.1515/jag-2015-0028.

6. Heinz E., Eling C., Wieland M., Klingbeil L., Kuhlmann H. Development, calibration and evaluation of a portable and direct georeferenced laser scanning system for kinematic 3D mapping // Journal of Applied Geodesy. – 2015. – № 9. – С. 227–243. – DOI 10.1515/jag-2015-0011.
7. Ramonell C., Chacon R. Open-source terrestrial laser scanner for the virtualization of geometrical entities in AEC classrooms // Computer Applications in Engineering Education. – 2022. – № 30. – С. 1009–1021. – DOI 10.1002/cae.22499. – EDN KICVTO.
8. Tang P., Huber D., Akinci B., Lipman R., Lytle A. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques // Automation in Construction. – 2010. – № 19. – С. 829–843. – DOI 10.1016/j.autcon.2010.06.007. – EDN OEJZJD.
9. Yang L., Cheng J., Wang Q. Semi-automated generation of parametric BIM for steel structures based on terrestrial laser scanning data // Automation in Construction. – 2020. – № 112. – С. 1–17. – DOI 10.1016/j.autcon.2019.103037. – EDN SNSGGQ.
10. Лисин А. Валидация информационной модели с использованием плагина BIM Interoperability Tools : магистерская диссертация. – Екатеринбург, 2023. – С. 59–61.
11. Султанов Ш., Кукина А. Интероперабельность программного обеспечения при проектировании сложных геометрических форм в BIM : материалы научно-практической Всероссийской конференции – Екатеринбург, 2021. – С. 51–68.
12. Wenlong L., Xiaoping Z., Baoguo X. Application of constructing three-dimensional model using laser scanning technology // Applied Mechanics and Materials. – 2011. – № 94. – С. 86–89. – DOI 10.4028/www.scientific.net/AMM.94-96.86.
13. Pdfplumber Documentation [Электронный ресурс]. – URL: <https://github.com/jsvine/pdfplumber> (дата обращения 05.09.2024).
14. Pillow (PIL Fork) Documentation [Электронный ресурс]. – URL: <https://pillow.readthedocs.io/en/stable/> (дата обращения 05.09.2024).
15. Pytesseract Documentation [Электронный ресурс]. – URL: <https://pypi.org/project/pytesseract/> (дата обращения 05.09.2024).
16. Configparser Documentation [Электронный ресурс]. – URL: <https://docs.python.org/3/library/configparser.html> (дата обращения 05.09.2024).
17. Openpyxl Documentation [Электронный ресурс]. – URL: <https://openpyxl.readthedocs.io/en/stable/> (дата обращения 05.09.2024).
18. Tkinter Documentation [Электронный ресурс]. – URL: <https://docs.python.org/3/library/tkinter.html> (дата обращения 05.09.2024).

Об авторах

Диас Мейрамович Искаков – аспирант, инженер-геодезист.

Получено 20.12.2024

© Д. М. Искаков, 2025

Automation of BIM modeling processes through project documentation

*D. M. Iskakov*¹✉

¹ Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation

e-mail: d.iskakov@svarog.ru

Abstract. The automation of the project documentation data processing and its integration into BIM models is considered in the paper. Particular attention is given to the data preparation phase that is conventionally regarded as the most time-consuming and resource-intensive step. The primary objec-

tive of the research is to achieve automation in BIM modeling through project documentation. The main tasks of the study are to automate data extraction from PDF drawings, their subsequent processing and conversion into a format suitable for use in BIM systems. The obtained results demonstrate that the developed tool can significantly reduce data processing time and improve the accuracy of their integration. The key conclusions drawn from the study indicate that the proposed tool simplifies and accelerates such processes as converting data from technical documentation, extracting textual and tabular information, verifying completeness and consistency, thereby enhancing the efficiency of creating component catalogs and mini-catalogs for pipeline systems, which contributes to increased design accuracy and reduced project timelines. This is particularly crucial when dealing with large volumes of data, where traditional methods may prove laborious and less accurate.

Keywords: Python, optical character recognition, process automation, PDF drawings, BIM, data mining, data integration

REFERENCES

1. Zaczek-Peplinska J., Pasik M., Adamek A., Kołakowska M., lapinski S. (2013). Monitoring technical conditions of engineering structures using the terrestrial laser scanning technology. *Reports on Geodesy and Geoinformatics*, 95, 1–10, DOI 10.2478/rgg-2013-0008.
2. Wang J., Kutterer H., Fang X. (2012). On the detection of systematic errors in terrestrial laser scanning data. *Journal of Applied Geodesy*, 6, 187–192, DOI 10.1515/jag-2012-0025.
3. Błaszczyk M., Laska M., Sivertsen A., Jawak S. (2022). Combined use of aerial photogrammetry and terrestrial laser scanning for detecting geomorphological changes in Hornsund, Svalbard. *Remote Sensing*, 14, 123–131, DOI 10.3390/rs14030601. EDN COHIBP.
4. Liu J., Fu L., Cheng G., Li D. (2022). Automated BIM reconstruction of full-scale complex tubular engineering structures using terrestrial laser scanning. *Remote Sensing*, 14, 542–553, DOI: 10.3390/rs14071659. EDN PDPXWC.
5. Dupuis J., Holst C., Kuhlmann H. (2016). Laser scanning based growth analysis of plants as a new challenge for deformation monitoring. *Journal of Applied Geodesy*, 10, 37–44, DOI: 10.1515/jag-2015-0028.
6. Heinz E., Eling C., Wieland M., Klingbeil L., Kuhlmann H. (2015). Development, calibration and evaluation of a portable and direct georeferenced laser scanning system for kinematic 3D mapping. *Journal of Applied Geodesy*, 9, 227–243. DOI 10.1515/jag-2015-0011.
7. Ramonell C., Chacon R. (2022). Open-source terrestrial laser scanner for the virtualization of geometrical entities in AEC classrooms. *Computer Applications in Engineering Education*, 30, 1009–1021, DOI 10.1002/cae.22499. EDN KICVTO.
8. Tang P., Huber D., Akinci B., Lipman R., Lytle A. (2010). Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in Construction*, 19, 829–843, DOI: 10.1016/j.autcon.2010.06.007. EDN OEJZJD.
9. Yang L., Cheng J., Wang Q. (2020). Semi-automated generation of parametric BIM for steel structures based on terrestrial laser scanning data. *Automation in Construction*, 112, 1–17, DOI: 10.1016/j.autcon.2019.103037. EDN SNSGGQ.
10. Lisin A. (2023). Validation of the information model using the BIM Interoperability Tools plug-in [Master's thesis]. Yekaterinburg, 59–61. [in Russian].
11. Sultanov Sh., Kukina A. (2021). Interoperability of software for designing complex geometric shapes in BIM [Proceedings of the Scientific and Practical All-Russian Conference]. Yekaterinburg, 51–68. [in Russian].
12. Wenlong L., Xiaoping Z., Baoguo X. (2011). Application of constructing three-dimensional model using laser scanning technology. *Applied Mechanics and Materials*, 94, 86–89, DOI: 10.4028/www.scientific.net/AMM.94-96.86.

13. Pdfplumber Documentation. Retrieved from <https://github.com/jsvine/pdfplumber> (accessed September 5, 2024).

14. Pillow (PIL Fork) Documentation. Retrieved from <https://pillow.readthedocs.io/en/stable/> (accessed September 5, 2024).

15. Pytesseract Documentation. Retrieved from <https://pypi.org/project/pytesseract/> (accessed September 5, 2024).

16. Configparser Documentation. Retrieved from <https://docs.python.org/3/library/configparser.html> (accessed September 5, 2024).

17. Openpyxl Documentation. Retrieved from <https://openpyxl.readthedocs.io/en/stable/> (accessed September 5, 2024).

18. Tkinter Documentation. Retrieved from <https://docs.python.org/3/library/tkinter.html> (accessed September 5, 2024).

Author details

Dias M. Iskakov – Ph. D. Student, surveyor engineer.

Received 20.12.2024

© *D. M. Iskakov*, 2025