

004.942

Формалізація моделей Інтернет-порталів знань

Л. С. Глоба, Н. В. Дерманська, Р. Л. Новогрудська

*Національний технічний університет України «Київський політехнічний інститут»,
Україна*

У статті розглянуто застосування системного підходу до побудови Інтернет-порталів знань. На базі системного підходу запропоновано моделі, необхідні для проектування Інтернет-порталів знань, проведена формалізація функціональної, структурної та інформаційної моделей порталу. Описано результати використання системного підходу для проектування Інтернет-порталу Національного Антарктичного Наукового Центру.

Ключові слова: *Інтернет-портал знань, системне проектування, системний підхід, формалізація, моделі.*

В статье рассмотрено использование системного подхода для построения Интернет-порталов знаний. На базе системного подхода предложены модели, необходимые для проектирования Интернет-порталов знаний, проведена формализация функциональной, структурной и информационной моделей портала. Описаны результаты использования системного подхода при проектировании Интернет-портала Национального Антарктического Научного Центра.

Ключевые слова: *Интернет-портал знаний, системное проектирование, системный подход, формализация, модели.*

The article deals with the usage of the system approach to designing knowledge Internet-portals. The models required for designing knowledge Internet-portals have been proposed on the base of the system approach. Formalization for functional, structural and information models is suggested. The results of the system approach usage for designing the Internet-portal of National Antarctic Scientific Center are described.

Key words: *knowledge Internet-portal, system engineering, systems approach, formalization, models.*

1. Системний підхід до проектування Інтернет-порталів знань

Інтернет-портал є веб-орієнтованим складним комплексом прикладного програмного забезпечення, який оперує одним або декількома джерелами інформації, обробляє її та подає інформацію заданої предметної області у зручному для користувача вигляді [1]. Портал знань є спеціалізованою інформаційною системою, що дозволяє здійснювати представлення, зберігання, систематизацію та структурування знань заданої предметної області, а також застосовувати до них ефективні методи пошуку [2].

У середовищі порталу знань зберігається велика кількість інтерактивних Інтернет-сервісів. Портал знань характеризується складною структурою та великою кількістю як інформаційних, так і функціональних елементів. Саме тому портал знань є складною системою та для його проектування застосовують методи системного проектування.

Методологія системного проектування дозволяє створювати метаописи (структуру, функції, інформаційні потоки та ін.) порталу з необхідним рівнем деталізації таких атрибутів:

- об'єкту проектування;
- моделі об'єкту (формалізоване представлення об'єкту);
- процесу проектування (методи, засоби, інструменти);
- результати цього процесу (технічна документація, що забезпечує організацію та здійснення всіх етапів).

Методологічною основою системного проектування є системний підхід – напрямок методології аналізу, синтезу і дослідження, основою якого є розгляд складного об'єкту як цілісної множини елементів у сукупності відношень та зв'язків між ними.

Одним з основним прийомів виконання системного підходу є декомпозиція процесу отримання рішення на послідовні логічно пов'язані етапи, на кожному з яких використовуються різноманітні формальні методи, моделі, інструментальні системи [3].

Використання системного підходу під час проектування Інтернет-порталів знань дозволяє структурувати та систематизувати інформацію, що зосереджена у середовищі порталу, а також створити процеси, які описують взаємодію потоків такої інформації. При використанні системного підходу в процесі проектування необхідно розробити функціональну, структурну та інформаційну моделі порталу.

Структурна модель описує компоненти, з яких складається створювана складна система, зв'язки між цими компонентами та принципи їх взаємодії. Для побудови структурної моделі можна застосовувати деревоподібну структуру як спосіб представлення ієрархічної структури у графічному вигляді [4].

Функціональна модель дозволяє описати повноту функцій системи. Для побудови функціональної моделі використовується стандарт Integration Definition for Function Modeling (IDEFO) [5].

Інформаційна модель описує інформаційні аспекти системи, яка моделюється. Усі об'єкти цієї моделі можуть бути представлені за допомогою креслень, схем, таблиць, діаграм, текстів. Для представлення інформаційної моделі застосовується діаграма Data Flow Diagram (DFD) за нотацією Гейна-Сарсона [6].

2. Формалізований опис моделей Інтернет-порталів знань

За допомогою формалізованого опису моделей Інтернет-порталу знань в роботі запропоновано уніфіковані формальні структури для представлення усіх елементів таких моделей. Такий формалізований опис може надалі бути використаний для оптимізації зв'язків та залежностей між елементами порталу.

Для опису Інтернет-порталу знань використано функціональну, структурну та інформаційну моделі, Система моделей порталу може бути представлена формалізмом:

$$S_M(P) = \{M_{IDEFO}, M_I, M_{DFD}\}, \quad (1)$$

де M_{IDEFO} – функціональна модель;

M_I – структурна модель;

M_{DFD} – інформаційна модель.

2.1 Формалізований опис функціональної моделі Інтернет-порталу знань

Задаємо функціональну модель таким формалізмом:

$$M_{IDEF0} = \langle B, D^{in}, D^{out}, C, W, L_i \rangle, \quad (2)$$

де $B = \{b_i, i = \overline{1, n}, n \in Z\}$ – множина блоків, кожний з яких відповідає певній функції;

$D^{in} = \{d_i^{in}, i = \overline{1, n}, n \in Z\}$ – множина вхідних даних;

$D^{out} = \{d_i^{out}, i = \overline{1, n}, n \in Z\}$ – множина вихідних даних;

$C = \{c_i, i = \overline{1, n}, n \in Z\}$ – множина елементів керування;

$W = \{w_i, i = \overline{1, n}, n \in Z\}$ – множина механізмів здійснення;

L_i – множина рівнів декомпозиції.

Рівень декомпозиції функціональної моделі визначає нумерацію блоків функцій:

$$B_i = \bigcup_{j=1}^m b_{ij}, i = \overline{0, n}, n \in Z, m \in Z. \quad (3)$$

За стандартом IDEF0 нумерація блоків має розпочинатися з «А», тому має місце відповідність:

$$b_{ij} \sim A_{pj}, \quad (4)$$

де p – порядковий номер блоку на рівні $L_i - 1$ для $L_i \geq 1$.

Для представлення відповідності елементу моделі його формалізованому позначенню та позначенню за стандартом IDEF0 використаємо табл. 1.

Табл. 1. Відповідність елементів функціональної моделі їх формалізованим позначенням та позначенням за стандартом IDEF0

Формалізоване позначення елементу функціональної моделі	Елемент функціональної моделі	Структурний елемент стандарту IDEF0
b_i	Блок	Функціональний блок
d_i^{in}	Вхідні дані	Вхід \longrightarrow
d_i^{out}	Вихідні дані	\longrightarrow Вихід
c_i	Керування	Керування \downarrow
w_i	Механізм	Механізм \uparrow

2.2 Формалізований опис структурної моделі Інтернет-порталу знань

Структурна модель визначається за допомогою ієрархічних рівнів та структурних елементів.

Кількість ієрархічних рівнів структурної моделі позначається N . Розрахунок номеру рівня починається з кореневого вузла, для якого $N=1$.

Кожен структурний елемент (СЕ) має позначення V_{ij} , де i відповідає ієрархічному рівні, на якому знаходиться СЕ, а j – номеру СЕ на цьому рівні.

Зобразимо відповідність елементів моделі їх позначенням для формалізованого опису та позначенням на дереві ієрархій (табл.2).

Табл.2. Відповідність елементів структурної моделі їх позначенням

Формалізоване позначення елементу структурної моделі	Елемент структурної моделі	Структурний елемент ієрархічної системи
V_{ij}	Вузол (СЕ)	Назва вузла
N	Кількість ієрархічних рівнів	Відсутній

2.3 Формалізований опис інформаційної моделі Інтернет-порталу знань

Інформаційна модель зображується за допомогою діаграми DFD, тому її формалізований вигляд визначається складовими DFD:

$$M_{DFD} = \langle E^{out}, S, R, P, T \rangle, \quad (5)$$

де $E^{out} = \{e_i^{out}, i = \overline{1, n}, n \in Z\}$ – множина зовнішніх сутностей;

$S = \{s_i, i = \overline{1, n}, n \in Z\}$ – множина систем і підсистем;

$R = \{r_i, i = \overline{1, n}, n \in Z\}$ – множина накопичувачів даних;

$P = \{p_i, i = \overline{1, n}, n \in Z\}$ – множина процесів;

$T = \{t_i, i = \overline{1, n}, n \in Z\}$ – множина потоків даних.

Покажемо відповідність елементів моделі їх позначенням (табл.3).

Табл.3. Відповідність елементів інформаційної моделі їх формалізованим позначенням та позначенням за стандартом DFD

Формалізоване позначення елементу інформаційної моделі	Елемент інформаційної моделі	Структурний елемент стандарту DFD
e_i^{out}	Зовнішня сутність	Ім'я зовнішньої сутності

s_i	Система або підсистема	<div style="border: 1px solid black; border-radius: 10px; padding: 5px;"> <div style="border-bottom: 1px solid black; padding-bottom: 2px;">Поле номеру підсистеми</div> <div style="border-bottom: 1px solid black; padding-bottom: 2px;">Поле імені підсистеми</div> <div style="padding-bottom: 2px;">Поле імені проектувальника підсистеми</div> </div>
r_i	Накопичувач	<div style="border: 1px solid black; padding: 5px;"> <div style="border-bottom: 1px solid black; padding-bottom: 2px;">Поле імені накопичувача</div> <div style="padding-bottom: 2px;">Поле номера накопичувача</div> </div>
p_i	Процес	<div style="border: 1px solid black; border-radius: 10px; padding: 5px;"> <div style="border-bottom: 1px solid black; padding-bottom: 2px;">Поле номеру процесу</div> <div style="border-bottom: 1px solid black; padding-bottom: 2px;">Поле імені процесу</div> <div style="padding-bottom: 2px;">Поле імені фізичної реалізації процесу</div> </div>
t_i	Потік	Потік даних \rightarrow

Наведений формалізований опис вводить формальне позначення для всіх елементів приведених моделей, що дозволяє задати уніфіковані контекстно-незалежні структури для їх представлення. Подана у таблицях відповідність між формалізованим описом елементів моделей та позначенням їх структурних елементів дозволяє полегшити процес автоматизації побудови самих моделей.

3. Проектування порталу Національного антарктичного наукового центру

В роботі проведено проектування порталу Національного Антарктичного Наукового Центру (НАНЦ) із застосуванням системного підходу та запропонованої системи моделей порталу [7]. Система моделей порталу НАНЦ включає: функціональну, структурну та інформаційну моделі. Для представлення кожної моделі побудовані діаграми з використанням відповідних нотацій. На базі запропонованих моделей проведена структуризація інформації, описано функціональні можливості порталу, а також, визначено обсяги потоків даних та відповідні workflow під час його функціонування.

Організація порталу НАНЦ передбачає забезпечення збору, обробки, передачі та представлення даних результатів антарктичних досліджень, для чого необхідно:

- створити єдине інформаційне середовище даних антарктичних досліджень України,
- реалізувати зберігання не лише даних, але й знань, які здобуті в результаті досліджень в Антарктиді;
- надати доступ до інформації дослідникам, які перебувають в експедиції;
- вдосконалити механізм систематизації та класифікації інформації;
- систематизувати дослідження для організації ефективного пошуку даних різних категорій та предметних областей;

- організувати ефективний пошук не лише за ключовими словами, але й у знайомих користувачу термінах предметної області [8,9].

Розглянемо опис інформаційної моделі порталу знань із застосуванням діаграми DFD (Рис. 1) на прикладі опису інформаційних зв'язків на порталі НАНЦ, потоків даних та перебігу інформації, що відбувається в процесі функціонування порталу [10].



Рис. 1 Інформаційна модель порталу НАНЦ

Формалізований опис даної моделі.

Користувачі порталу представлені як множина E^{out} :

e_1^{out} - 1-ий користувач;

...

e_k^{out} - k -ий користувач.

Для порталу НАНЦ множина зовнішніх сутностей представлена як:

$E^{out} = \{e_1^{out}, \dots, e_k^{out}, k \in Z\}$, де індекс k відповідає кількості користувачів.

До складу множини накопичувачів даних входять:

r_1 - база даних наукових знань, даних та результатів досліджень;

r_2 - спискові структури (сховища даних, ієрархічні структури файлів даних).

Таким чином, для порталу НАНЦ множина накопичувачів даних представлена як:

$$R = \{r_1, r_2\}.$$

Множина процесів включає такі елементи:

p_1 - «Шукати запитані дані»;

p_2 - «Переглянути»;

p_3 - «Заповнити».

Для порталу НАНЦ множина зовнішніх сутностей представлена як:

$$P = \{p_1, p_2, p_3\}.$$

Множина потоків даних представлена елементами:

t_1 - потік «Запити користувача», що йде від зовнішньої сутності «Користувач» до процесу «Шукати запитані дані»;

t_2 - потік «Запити користувача», що йде від зовнішньої сутності «Користувач» до процесу «Переглянути»;

t_3 - потік «Наукові знання, дані та результати досліджень, особиста інформація дослідника», який прямує від зовнішньої сутності «Користувач» до процесу «Заповнити»;

t_4 - потік «Результат запити», що з'єднує накопичувач «База даних наукових знань, даних та результатів досліджень» з процесом «Шукати запитані дані»;

t_5 - потік «Результат запити», що з'єднує накопичувач «База даних наукових знань, даних та результатів досліджень» з процесом «Переглянути»;

t_6 - потік «Результат запити», що з'єднує накопичувач «Спискові структури» з процесом «Шукати запитані дані»;

t_7 - потік «Результат запити», що з'єднує накопичувач «Спискові структури» з процесом «Переглянути»;

t_8 - потік «Внесення нової інформації», який створює зв'язок процесу «Заповнити» з накопичувачем «База даних наукових знань, даних та результатів досліджень»;

t_9 - потік «Внесення нової інформації», який створює зв'язок процесу «Заповнити» з накопичувачем «Спискові структури»;

t_{10} - потік «Наукові знання, дані та результати досліджень, нормативно-правові акти, шаблони досліджень, персональний кабінет», що поєднує процес «Шукати запитані дані» з зовнішньою сутністю «Користувач».

t_{11} - потік «Наукові знання, дані та результати досліджень, нормативно-правові акти, шаблони досліджень, персональний кабінет», що поєднує процес «Переглянути» з зовнішньою сутністю «Користувач».

Таким чином, для порталу НАНЦ множина потоків даних має наступний вигляд:

$$T = \{t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9, t_{10}, t_{11}\}.$$

Розглянемо опис функціональної моделі порталу знань за стандартом IDEF0 на прикладі опису функціональних можливостей порталу НАНЦ [10].

Розглянемо функціональну модель порталу НАНЦ після першої декомпозиції головної функції (Рис. 2).

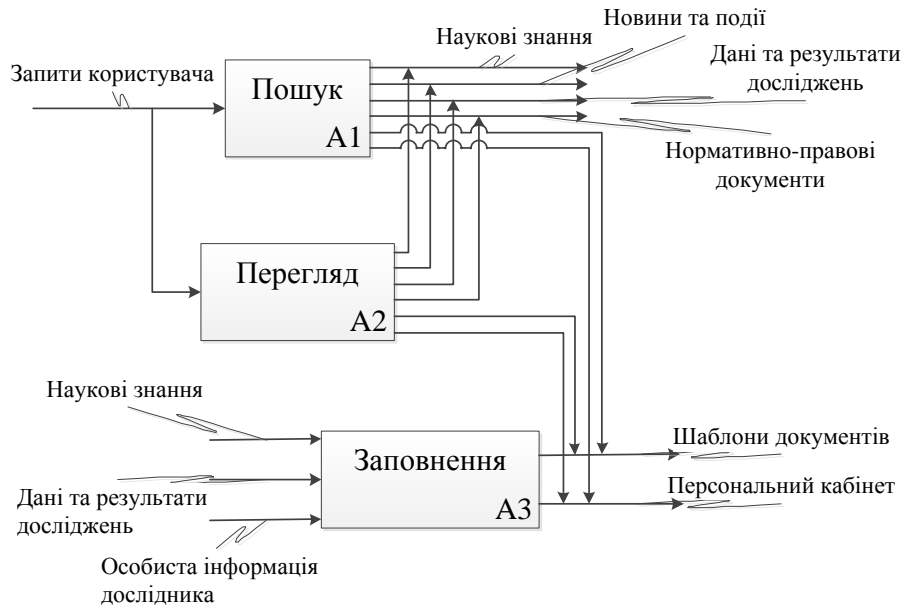


Рис. 2 Функціональна модель порталу НАНЦ після першої декомпозиції головної функції

Формалізований опис даної моделі.

Оскільки розглядається функціональна модель порталу після першої декомпозиції головної функції, то $L_i = 1$.

Множина функціональних можливостей порталу представлена як множина B , де:

b_1 - пошук;

b_2 - перегляд;

b_3 - заповнення.

Таким чином, для порталу НАНЦ множина доступних функцій представлена як:

$$B = \{b_1, b_2, b_3\}.$$

Множина вхідних даних включає такі елементи:

d_1^{in} - 1-ий запит користувача;

...

d_k^{in} - k -ий запит користувача;

d_{k+1}^{in} - наукові знання;

d_{k+2}^{in} - дані та результати досліджень;

d_{k+3}^{in} - особиста інформація дослідника.

Для порталу НАНЦ множина вхідних даних представлена як:

$D^{in} = \{d_1^{in}, \dots, d_{k+3}^{in}, k \in Z\}$, де індекс k відповідає кількості запитів користувача.

Множина вихідних даних представлена елементами:

Множина вихідних даних представлена елементами:

d_1^{out} - наукові знання;

d_2^{out} - новини та події;

d_3^{out} - дані та результати досліджень;

d_4^{out} - нормативно-правові документи;

d_5^{out} - шаблони документів;

d_6^{out} - персональний кабінет.

Таким чином, для порталу НАНЦ множина вихідних даних має наступний вигляд:

$D^{out} = \{d_1^{out}, d_2^{out}, d_3^{out}, d_4^{out}, d_5^{out}, d_6^{out}\}$.

Так як, множина елементів керування та множина механізмів здійснення не представлені на даній моделі, то:

$C = \emptyset$;

$W = \emptyset$.

Розглянемо функціональну модель порталу НАНЦ для декомпозиції функції заповнення (Рис. 3).

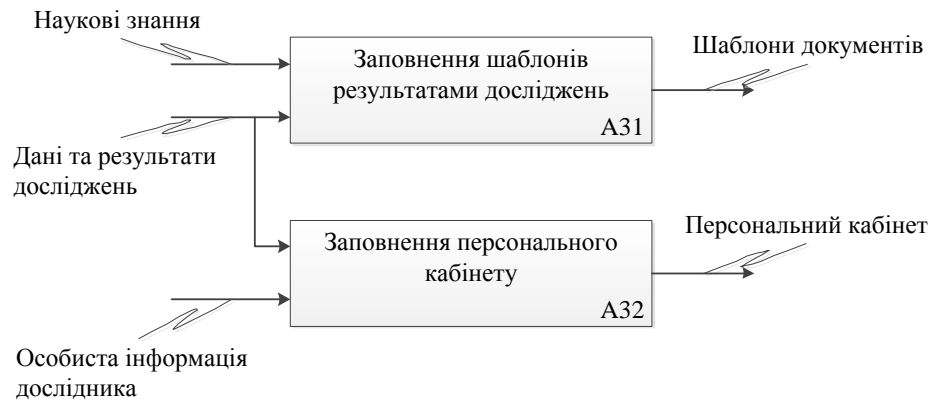


Рис. 3 Функціональна модель порталу НАНЦ для декомпозиції функції заповнення

Оскільки розглядається функціональна модель порталу після першої декомпозиції головної функції, то $L_i = 2$.

Множина функціональних можливостей порталу представлена як множина B_3 , де:

b_{31} - заповнення шаблонів результатами досліджень;

b_{32} - заповнення персонального кабінету.

Таким чином, для порталу НАНЦ множина доступних функцій представлена як:

$$B_3 = \{b_{31}, b_{32}\}.$$

Множина вхідних даних включає такі елементи:

d_1^{in} - наукові знання;

d_2^{in} - дані та результати досліджень;

d_3^{in} - особиста інформація дослідника.

Для порталу НАНЦ множина вхідних даних представлена як:

$$D^{in} = \{d_1^{in}, d_2^{in}, d_3^{in}\}.$$

Множина вихідних даних представлена елементами:

d_1^{out} - шаблони документів;

d_2^{out} - персональний кабінет.

Таким чином, для порталу НАНЦ множина вихідних даних має наступний вигляд:

$$D^{out} = \{d_1^{out}, d_2^{out}\}.$$

Так як, множина елементів керування та множина механізмів здійснення не представлені на даній моделі, то:

$$C = \emptyset;$$

$$W = \emptyset.$$

Застосування системного підходу до проектування порталу НАНЦ надало можливість підвищити ефективність опису досліджень проведених науковцями станції Академік Вернадський. Системне проектування дозволило систематизувати та класифікувати інформацію про проведені в Антарктиді дослідження науковцями, що працюють у різних галузях знань. Використання множини моделей створених на базі системного підходу надало можливість автоматизувати процеси розробки, підтримки функціонування та розвитку порталу знань, а також представити елементи цих моделей з засобами проектування інформаційних систем.

4. Висновки

У роботі описано використання системного підходу до проектування складних інформаційних систем. Показано, що Інтернет-портали знань є складною системою і для їх проектування доцільно застосовувати системний підхід.

Проведено аналіз видів моделей, які використовуються при проектуванні Інтернет-порталів знань, способи та нотації для їх представлення.

Запропонований формалізований опис узагальненої моделі Інтернет порталу знань. Проведено формалізацію елементів інформаційної, функціональної та структурної моделей, що дозволило задати уніфіковані контекстно-незалежні структури для їх представлення.

ЛІТЕРАТУРА

1. What is a web portal and what types of portals [Електронний ресурс]. – Режим доступу: <http://community.telecentre.org/profiles/blogs/what-is-a-web-portal-and-what>.
2. Концепция интеллектуального интернет-портала знаний для доступа к информационным ресурсам по археологии и этнографии [Електронний ресурс]. – Режим доступу: http://www.ssc.smr.ru/media/ipuss_conf/06/4_08.pdf.
3. Гольшев Л. К. Системный подход к анализу и проектированию сложных систем. Системный проектирование / Л. К. Гольшев: научн. моногр. – К. : ГП «Информационно-аналитическое агентство», 2011. – 555 с.
4. Системный анализ : проблемы, методология, приложения / Михаил Захарович Згуровский, Наталья Дмитриевна Панкратова, Институт прикладного системного анализа НАН Украины . – 2-е изд., перераб. и доп. – Київ : Наукова думка, 2011 . – 726 с.
5. Лавров В. В. Разработка функциональной модели автоматизированной информационной системы анализа и прогнозирования работы доменного цеха / В. В. Лавров, Н. А. Спиринов Николай, А. А. Бурькин и др. // Известия Томского политехнического университета. – №5, 2011. – С.137–143.
6. Методы моделирования и модели разработки ИС [Електронний ресурс]. – Режим доступу: <http://infdis.narod.ru/pis/pis-p3-1.htm>
7. Pryvar O.A., Globa L.S., Novogrudskaya R.L. Sozdanie Ukrainskogo portala Antarkticheskikh dannykh [Ukrainian Antarctic data portal developing]. 7-ya mezhdunarodnaja konferencyja “Problemy telekommunikacyj - 2013”, 2013.
8. Мороз І.В. Создание единого информационного пространства данных антарктических исследований / Мороз І.В., Глоба Л.С., Новогрудская Р.Л., Мочалкина К.С., Кузін І.О. // Український Антарктичний Журнал. – 2011. – № 10–11. – С. 343–351.
9. Щасливий С.П. Автоматизація робочого процесу обробки даних Національного Антарктичного / С.П.Щасливий, Л.С. Глоба, Р.Л. Новогрудська // Міжнародна науково-технічна конференція «Проблеми телекомунікацій»: Збірник тез. – К.: НТУУ «КПІ», 2014, – С. 267-269
10. Дерманська Н. В. Системний підхід до моделювання порталу «Національний центр антарктичних даних» / Н. В. Дерманська, Р. Л. Новогрудська // Український Антарктичний Журнал. – 2015. – № 14. – С. 238-245