

УДК 004.62

А.В. Меленець

Державний департамент страхового фонду документації МНС України, Харків

ОБРАННЯ ТИПУ СХОВИЩА ДЛЯ ЗБЕРІГАННЯ ДОКУМЕНТІВ СТРАХОВОГО ФОНДУ В ЕЛЕКТРОННОМУ ВИГЛЯДІ

Приведено структуру документа страхового фонду та його копії в електронному вигляді. Розглянуто різні варіанти зберігання документів страхового фонду в електронному вигляді, визначено критерії для вибору типу сховища. Методом аналізу ієрархій обрано тип сховища для зберігання документів страхового фонду в електронному вигляді. Запропоновано використання технології сховищ даних для зберігання документів страхового фонду в електронному вигляді.

Ключові слова: документ страхового фонду, база даних, сховище даних, реєстр, інтеграція баз даних, критерії, метод аналізу ієрархій, надзвичайна ситуація, аварія, потенційно небезпечний об'єкт.

Вступ

Прийняття будь-якого рішення ґрунтується перш за все на наявній інформації про об'єкт управління. У випадках, коли для прийняття рішень є досить невеликий проміжок часу, вага інформаційного забезпечення суттєво підвищується. На цей час у багатьох країнах світу ведуться роботи по створенню інтегрованих систем підтримки прийняття рішень під час ліквідації аварій та надзвичайних ситуацій. Такі інтегровані системи мають зберігати найрізноманітнішу інформацію про потенційно небезпечні об'єкти. Одним із великих масивів інформації про небезпечні об'єкти є страховий фонд документації. Страховий фонд документації (СФД) – упорядкований банк документів, зафіксованих на мікрографічній плівці чи інших компактних носіях інформації, які прийняті на державний облік і довгострокове надійне зберігання. На цей час в Україні до СФД закладається увесь необхідний комплект проектної документації для будівництва об'єктів. До такої документації відносяться генеральний план об'єкта, поповерхові плани будівель та споруд, схеми інженерних мереж та комунікацій тощо, яка може бути ефективно використана під час ліквідації аварій та надзвичайних ситуацій.

Згідно з існуючою світовою практикою, документи страхового фонду зберігаються на мікрофільмах. Мікрофільми знаходяться на спеціальних базах зберігання, де дотримуються суворі кліматичні вимоги і для відновлення документації з мікрофільму необхідний деякий час: до шести годин на акліматизацію мікрофільмів та саме відновлення, при якому виконується сканування мікрофільмів, доопрацювання отриманих кадрів, їх зшивка та перенесення на необхідний носій (папір або електронний носій інформації), витрати часу на виконання відновлення можуть складати від 10 хвилин на 1 аркуш. Усе це не дозволяє оперативно отримати необхідні креслення об'єктів, тому доцільним було б зберігати необхідні

креслення на найбільш небезпечні об'єкти в уже готовому до використання вигляді на електронних носіях, а враховуючи те, що на цей час в Україні нараховується понад 20 тис. потенційно небезпечних об'єктів (ПНО) постає питання щодо зберігання таких комплектів документації в електронному вигляді.

Проблема зберігання документів СФД або аналогічних даних в електронному вигляді розглядалася у наукових працях [5 – 7, 10]. Зокрема, у [5] розглянуто питання довгострокового та надійного збереження електронних документів бібліотек та архівів та оптичних дисках, у роботах [6, 7, 10] розглянуто питання зберігання корпоративних даних у реляційних базах даних та сховищах даних. Але питання зберігання документів СФД в електронному вигляді до цього часу не розглядалось.

Метою статті є вибір типів зберігання документів СФД в електронному вигляді. Стаття структурована наступним чином: перший розділ присвячений аналізу структури документації СФД в електронному вигляді, визначення можливих об'ємів такої інформації, у другому розділі визначено варіанти зберігання документів СФД в електронному вигляді та критерії обрання, у третьому розділі з використанням методу аналізу ієрархій проведено вибір варіанту зберігання та у четвертому розділі надано опис використання технології сховищ даних для зберігання документів СФД в електронному вигляді

1. Документи страхового фонду в електронному вигляді

В узагальненому вигляді мікрофільм страхового фонду має таку структуру: титульні кадри, відомість комплекту замікрофільмованої документації та сама документація [2]. На титульних кадрах наводиться загальна інформація про об'єкт та замікрофільмована документація, відомість комплекту замікрофільмованої документації (ВК) за своєю суттю є змістом або переліком замікрофільмованих документів, з зазна-

ченням їх реквізитів. При зйомці документів використовується розбивка великих форматів на менші за спеціальною схемою мікрофільмування [2].

Процес створення комплексу документів в електронному вигляді виглядає так: сканування кадрів мікрофільму на спеціальному сканері мікрофільмів, їх доопрацювання у графічному редакторі, збереження у необхідному растровому форматі, отримання ВК з Державного реєстру документів СФД та саме побудова комплексу документації, яка передбачає формування на базі ВК файлу з переліком гіперпосилань на збережені растрові зображення креслень (рис. 1) [3].

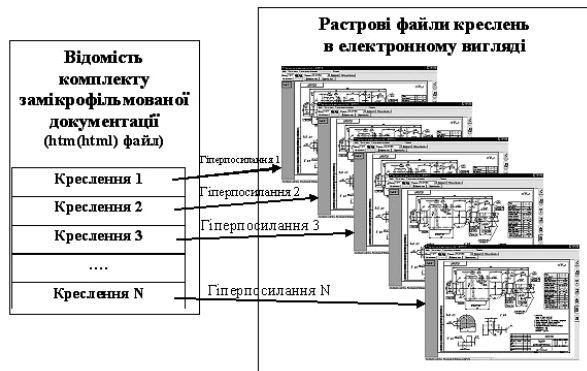


Рис. 1. Структура документа СФД в електронному вигляді

Визначення об'єму документів СФД в електронному вигляді. Оскільки об'єм файлу гіперпосилань у форматах htm (html) є незрівнянно малим у порівнянні з растровими зображеннями креслень, тому для розрахунку об'єму документів СФД в електронному вигляді використовуємо лише растрові зображення, які за вимогами нормативної бази [3] мають відповідати таким вимогам:

- роздільна здатність не менше 400 dpi;
- кольорова гама: градації сірого або чорно-біле зображення;
- формати TIFF (без стиснення), JPG;
- розмір зображення має відповідати оригіналу.

За результатами статистичних розрахунків у середньому інформація на 1 об'єкт складає приблизно 1000 аркушів формату А4. Розмір растрового файлу формату А4 виконаний відповідно до наведених вимог у форматі JPG від 2Mb.

Тобто можливо визначити, що об'єм документа СФД в електронному вигляді буде складати від 2Gb.

Враховуючи те, що на цей час у Державному реєстрі ПНО зареєстровано понад 20 тис. небезпечних об'єктів, об'єм бази документів СФД в електронному вигляді для цих небезпечних об'єктів може сягати до 40Tb.

Державний реєстр документів СФД. Державний реєстр документів СФД – це автоматизована система з обмеженим доступом, яка призначена забезпечити накопичення, облік, опрацювання та пошук документів страхового фонду.

На цей час Державний реєстр документів СФД містить усю інформацію про замикрофільмовану документацію, об'єкт, документація на який замикрофільмована, відомість комплексу замикрофільмованої документації, дані про мікрофільм та його характеристики та інші. Однак за своїм призначенням Державний реєстр документів СФД являє собою оперативне джерело даних і це потрібно враховувати при визначенні типу зберігання документів СФД в електронному вигляді.

2. Обрання варіанту зберігання документів СФД в електронному вигляді

2.1. Варіанти та критерії зберігання документів СФД в електронному вигляді. Для зберігання документів СФД в електронному вигляді можна виділити такі варіанти, як:

1. Збереження файлу гіперпосилань та растрових файлів на окремих DVD або Blu-ray носіях – варіант V1.
2. Збереження у базі даних Державного реєстру документів СФД – варіант V2.
3. Збереження у окремій реляційній базі даних – варіант V3.
4. Збереження у сховищі даних – варіант V4.

Для обрання варіанту носія для збереження документів СФД в електронному вигляді можна встановити такі критерії: K1 – швидкість отримання інформації; K2 – надійність збереження інформації; K3 – дотримання вимог технічного захисту інформації; K4 – можливість інтеграції з іншими інформаційними системами з питань надзвичайних ситуацій.

Оскільки запропоновані критерії є специфічними і для них досить складно визначити чіткі кількісні оцінки, то для прийняття рішення щодо вибору типу збереження документів СФД в електронному вигляді будемо використовувати метод аналізу ієрархій [4], який дозволить отримати результат шляхом попарного порівнювання ступеня реалізації критеріїв при використанні різних типів інтеграції.

2.2. Збереження на окремих DVD або Blu-ray носіях. Таке зберігання дозволяє досить легко вирішити питання щодо технічного захисту інформації, який можливо реалізувати лише організаційними заходами. Однак існують напрями, у яких використання оптичних дисків є проблематичним:

- по-перше, відносно надійності збереження даних на оптичних носіях; проведені дослідження щодо прискореного старіння оптичних дисків [5] довели, що вже через декілька років є можливість втрати інформації та і самого носія;
- по-друге, за наявності великого масиву оптичних дисків, що зберігаються окремо, швидкість отримання даних з диска незрівнянно нижча ніж у порівнянні з базою даних;
- по-третє, складність інтеграції інформації, яка зберігається на масиві оптичних дисків з інформаційними системами.

2.3. Збереження у базі даних Державного реєстру документів СФД. Для такого типу зберігання документів СФД в електронному вигляді можна відзначити те, що питання технічного захисту інформації у Державному реєстрі документів СФД вирішені, однак слід відзначити, що обсяг документів СФД в електронному вигляді може складати терабайти і збереження такої інформації у Державному реєстрі документів СФД суттєво збільшить обсяг його бази даних та значно знизить його продуктивність, збільшить час відклику на запити із інших систем. Крім того склад таких даних є переважно статичним з низькою частотою використання, що призведе до наявності у базі даних областей, які практично не використовуються.

2.4. Збереження у окремій реляційній базі даних та у сховищі даних. Для такого типу зберігання необхідно створення нової інформаційної системи, одночасно необхідно вирішувати питання технічного захисту інформації та інші, при цьому система може бути побудована з урахуванням необхідності інтеграції з іншими системами. Однак слід відзначити, що склад даних документів СФД є переважно статичним зі специфічною структурою і використання для побудови реляційної бази даних призведе до зменшення продуктивності такої системи. Вирішити це питання можливо шляхом застосування технології сховищ даних [6]. При цьому розглядається варіант фізичного сховища або вітрини даних [7].

3. Обрання варіанту зберігання документів СФД в електронному вигляді

3.1. Порівняння критеріїв вибору. Відповідно до визначеної у методі аналізу ієрархій [4] шкалою, на підставі результатів попарного порівняння важливості критеріїв визначимо ступінь їх впливу (пріоритет). Для цього будується матриця А (1) попарних порівнянь критеріїв K_i та K_j за ступенем їх важливості:

$$A = \begin{bmatrix} 1 & 0.333333 & 0.5 & 0.333333 \\ 3 & 1 & 2 & 0.5 \\ 2 & 0.5 & 1 & 0.5 \\ 3 & 2 & 2 & 1 \end{bmatrix} \quad (1)$$

Виконуємо розрахунок вектору пріоритетів В (2) критеріїв K_1 - K_4 по даній матриці шляхом визначення головного власного вектору, який після нормалізації стає вектором пріоритетів [7], при цьому V_i – числове значення пріоритету K_i критерію:

$$B = \begin{bmatrix} 0.107958 \\ 0.293031 \\ 0.18719 \\ 0.41182 \end{bmatrix} \quad (2)$$

3.2. Порівняння за швидкістю отримання інформації. За результатами попарного порівняння кожного варіанту зберігання V_i за критерієм K_1 отримуємо матрицю С (3), при цьому числове значення C_{ij} відповідає рівню реалізації критерію K_1 при використанні варіанту зберігання V_i у порівнянні з V_j :

чення C_{ij} відповідає рівню реалізації критерію K_1 при використанні варіанту зберігання V_i у порівнянні з V_j :

$$C = \begin{bmatrix} 1 & 0.333333 & 0.2 & 0.142857 \\ & 1 & 0.5 & 0.2 \\ 5 & 2 & 1 & 0.5 \\ 7 & 5 & 2 & 1 \end{bmatrix} \quad (3)$$

Головне власне значення матриці С (3) – $\lambda_{\max} = 4.103964$.

Виконуємо розрахунок вектору пріоритетів D (4) матриці С (3). При цьому D_i – числове значення, яке відповідає ступеню реалізації критерію K_1 при використанні V_i варіанту зберігання:

$$D = \begin{bmatrix} 0.058518 \\ 0.137791 \\ 0.273522 \\ 0.530169 \end{bmatrix} \quad (4)$$

3.3. Порівняння за надійністю збереження інформації. За результатами попарного порівняння кожного варіанту зберігання V_i за критерієм K_2 отримуємо матрицю E (5), при цьому E_{ij} – кількісне значення рівню реалізації критерію K_2 при використанні варіанту зберігання V_i у порівнянні з V_j .

$$E = \begin{bmatrix} 1 & 0.2 & 0.2 & 0.2 \\ 5 & 1 & 1 & 0.5 \\ 5 & 1 & 1 & 0.5 \\ 5 & 2 & 2 & 1 \end{bmatrix} \quad (5)$$

Головне власне значення матриці E (5) – $\lambda_{\max} = 4.086546$.

Виконуємо розрахунок вектору пріоритетів F (6) матриці E (5). При цьому F_i – числове значення, яке відповідає ступеню реалізації критерію K_2 при використанні V_i варіанту зберігання:

$$F = \begin{bmatrix} 0.062162 \\ 0.253991 \\ 0.253991 \\ 0.429857 \end{bmatrix} \quad (6)$$

3.4. Порівняння за дотриманням вимог технічного захисту інформації. За результатами попарного порівняння кожного варіанту зберігання V_i за критерієм K_3 отримуємо матрицю G (7), при цьому G_{ij} – кількісне значення рівню реалізації критерію K_3 при використанні варіанту зберігання V_i у порівнянні з V_j :

$$G = \begin{bmatrix} 1 & 5 & 7 & 7 \\ 0.2 & 1 & 3 & 3 \\ 0.142857 & 0.333333 & 1 & 3 \\ 0.142857 & 0.333333 & 0.333333 & 1 \end{bmatrix} \quad (7)$$

Головне власне значення матриці G (7) – $\lambda_{\max} = 4.027758$.

Виконуємо розрахунок вектору пріоритетів H (8) матриці G (7). При цьому H_i – числове значення, яке відповідає ступеню реалізації критерію K_3 при використанні V_i варіанту зберігання:

$$H = \begin{bmatrix} 0.635181 \\ 0.190902 \\ 0.112169 \\ 0.061749 \end{bmatrix}. \quad (8)$$

3.5. Порівняння щодо простоти інтеграції з іншими системами. За результатами попарного порівняння кожного варіанту зберігання V_i за критерієм K_4 отримуємо матрицю L (9), при цьому L_{ij} – кількісне значення рівню реалізації критерію K_4 при використанні варіанту зберігання V_i у порівнянні з V_j :

$$L = \begin{bmatrix} 1 & 0.333333 & 0.2 & 0.2 \\ 3 & 1 & 0.333333 & 0.333333 \\ 5 & 3 & 1 & 1 \\ 5 & 3 & 1 & 1 \end{bmatrix}. \quad (9)$$

Головне власне значення матриці L (9) – $\lambda_{\max} = 4.066784$.

Виконуємо розрахунок вектору пріоритетів M (10) матриці L (9). При цьому M_i – числове значення, яке відповідає ступеню реалізації критерію K_4 при використанні V_i варіанту зберігання.

$$M = \begin{bmatrix} 0.068694 \\ 0.153452 \\ 0.388927 \\ 0.388927 \end{bmatrix}. \quad (10)$$

3.6. Результат порівняння варіантів зберігання. Для отримання результатів порівняння побудуємо на основі векторів пріоритетів D (4), F (6), H (8) та M (10) з урахуванням ваги кожного критерію (вектор пріоритетів B (2)) нову матрицю N (11) [7]. При цьому: $N_{ij} = D_i \cdot B_j$:

$$N = \begin{bmatrix} 0.02233 & 0.07442 & 0.4799 & 0.051901 \\ 0.05258 & 0.304077 & 0.144233 & 0.115938 \\ 0.104375 & 0.304077 & 0.084747 & 0.293847 \\ 0.20231 & 0.514623 & 0.046653 & 0.293847 \end{bmatrix}. \quad (11)$$

Виконуємо розрахунок вектору пріоритетів O (12) матриці N (11). При цьому O_i – числове значення, яке відповідає ступеню реалізації усіх критеріїв K_1 - K_4 при використанні V_i варіанту зберігання:

$$O = \begin{bmatrix} 0.206139 \\ 0.184034 \\ 0.257152 \\ 0.352675 \end{bmatrix}. \quad (12)$$

Найбільше значення у отриманому векторі пріоритетів O (12) має елемент O_4 .

Таким чином, найбільш прийнятним типом зберігання документів СФД в електронному вигляді є сховище даних.

4. Використання технології сховищ даних для зберігання документів СФД в електронному вигляді

Під час організації зберігання документів СФД в електронному вигляді виникає питання щодо зме-

ншення витрат часу та трудомісткості робіт із завантаження отриманих растрових зображень у сховище. На цей час Державний реєстр документів СФД містить великий обсяг структурованої інформації про об'єкт, документація на який замікрофільмована, до таких даних слід віднести: назва об'єкта, реквізити власника, коди за Класифікатором адміністративно-територіального устрою України, Єдиного державного реєстру підприємств, установ та організацій тощо, назва комплексу документації, ВК тощо.

Враховуючи вищевикладене, необхідним є не лише завантаження до сховища документів СФД в електронному вигляді але і відповідних даних про об'єкт із Державного реєстру документів СФД, з метою недопущення їх повторного набору та наявності зв'язки з базою даних Державного реєстру документів СФД. Тому більш доцільним є внесення документів СФД в електронному вигляді до бази даних Державного реєстру документів СФД, прив'язки до іншої детальної інформації та подальше перенесення документації в електронному вигляді та копіювання детальних даних до сховища документів СФД в електронному вигляді з використанням спеціального диспетчера завантаження (рис. 2). Після чого з бази даних Державного реєстру документів СФД слід вилучати документи в електронному вигляді та додавати посилання на сховище документів СФД.

Враховуючи вищевикладене та сучасний стан уявлення архітектури сховищ даних [6, 7, 10], загальна архітектура сховища даних документів страхового фонду в електронному вигляді має складатись з таких компонентів (рис. 3):

- детальні дані (характеристики об'єкта, відомість комплексу замікрофільмованої документації та креслення документів страхового фонду), які поступають з Державного реєстру документів СФД;
- метадані, які мають описувати джерело даних і будь-які зміни, внесені в ці дані, та спосіб зберігання даних у сховищі;
- агреговані дані, у якості яких має виступати побудований відповідно до відомості комплексу замікрофільмованої документації комплект документів на об'єкт;
- диспетчер завантаження, який вилучає дані із Державного реєстру документів СФД та завантажує їх у тимчасову область зберігання у сховищі даних;
- диспетчер сховища, який виконує аналіз несуперечливості даних, перетворення й переміщення вихідних даних з тимчасової області зберігання в основні таблиці сховища даних, формує індекси тощо, виконує агрегування даних та інші операції;
- диспетчер запитів, який виконує усі операції, пов'язані з керуванням запитів від користувачів.

Використання такої архітектури сховища дозволить інтегрувати його до інших автоматизованих систем, у такому випадку диспетчер запитів буде виконувати роль інтеграційної шини при викорис-

тання інтеграції на основі єдиної понятійної моделі предметної області (concept-centric) [1].



Рис. 2. Схема завантаження даних у сховище

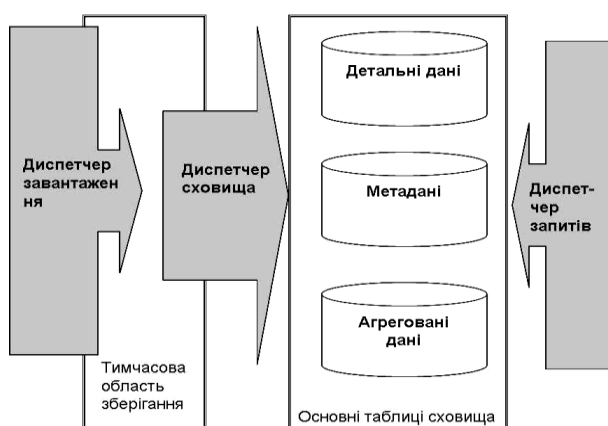


Рис. 3. Архітектура сховища документів СФД в електронному вигляді

Висновки

У статті проведено обґрунтування зберігання документів СФД в електронному вигляді у сховищі даних. Запропонована до використання технологія сховищ даних дозволить швидко отримувати інформацію, забезпечить надійність її збереження, дозволить додержати вимоги щодо технічного захисту інформації і призведе до інтеграції страхового

фонду документації з інформаційними системами з питань надзвичайних ситуацій.

Напрями подальшого розвитку у побудові сховища документів СФД в електронному вигляді пов'язані із розробленням архітектури самого сховища, обранням необхідних інструментів роботи з даними, визначення технології сховища, засобів інтеграції сховища з іншими інформаційними системами тощо.

Список літератури

1. Приемы объектно-ориентированного проектирования Паттерны Проектирования / Э. Гамма, Р. Хелм, Р. Джонсон, Дж. Влссидес. – СПб: Питер, 2001. – 368 с.
2. ДСТУ В 33/101-97 "СФД. Порядок побудови мікрофільмів страхового фонду".
3. ДСТУ 33.116:2010 "СФД. Електронні копії документів. Загальні вимоги".
4. Саати Т. Принятие решений. Метод анализа иерархий / Т. Саати. – М.: Радио и связь, 1993. – 278 с.
5. Тихонова И. К проблеме хранения CD и DVD дисков [Электронный ресурс] / И. Тихонова // 11-я ежегодная межд. Конф. «EVA 2008 Москва» – Режим доступа к ресурсу: www.evarussia.ru/upload/doklad/doklad_911.doc.
6. Коннолли Т. Базы данных. Проектирование, реализация сопровождение. Теория и практика / Т. Коннолли, К. Бег. М.: Изд. дом "Вильямс", 2003. – 1440 с.
7. Методы и модели анализа данных: OLAP и Data Mining / А. Баргесян, М. Куприянов, В. Степаненко, И. Холод. – СПб: БХВ-Петербург, 2004. – 336 с.
8. Аляев Ю. Дискретная математика и математическая логика / Ю. Аляев, С. М. Тюрин. – М.: Финансы и статистика, 2006. – 365 с.
9. Лиффенгуэлл Д. Принципы работы с требованиями к программному обеспечению. Унифицированный подход / Д. Лиффенгуэлл, Д. Уидриг. – М.: Издательский дом "Вильямс", 2002. – 445 с.
10. Хиббс Л. Oracle9iR2: разработка и эксплуатация хранилищ баз данных / Л. Хиббс, С. Хилсон, Ш. Лоуенд. – М.: КУДИЦ-ОБРАЗ, 2004. – 592 с.

Надійшла до редколегії 23.09.2010

Рецензент: д-р техн. наук, проф. В.С. Харченко, Національний аерокосмічний університет ім. М.Є. Жуковського «ХАІ», Харків.

ВЫБОР ТИПА ХРАНИЛИЩА ДЛЯ ХРАНЕНИЯ ДОКУМЕНТОВ СТРАХОВОГО ФОНДА В ЭЛЕКТРОННОМ ВИДЕ

А.В. Меленец

Приведена структура документа страхового фонда и его копии в электронном виде. Рассмотрены различные варианты хранения документов страхового фонда в электронном виде, определены критерии для выбора типа хранилища. Методом анализа иерархий выбран тип хранилища для хранения документов страхового фонда в электронном виде. Предложено использование технологии хранилищ данных для хранения документов страхового фонда в электронном виде.

Ключевые слова: документ страхового фонда, база данных, хранилище данных, реестр, интеграция баз данных, критерий, метод анализа иерархий, чрезвычайная ситуация, авария, потенциально опасный объект.

CHOICE OF TYPE OF STOREHOUSE FOR STORAGE OF DOCUMENTS OF INSURANCE FUND IN ELECTRONIC FORM

A.V. Melenets

The structure of the insurance fund document (IFD) and its copies in electronic form are resulted. Various variants of storage of IFDs in electronic form are considered, criteria for a choice of type of storehouse are defined. The method of the analysis of hierarchies was used to choose type of storehouse for storage of IFDs in electronic form. Use of technology of data warehouse for storage of IFDs in electronic form is offered.

Keywords: document of insurance fund, database, data warehouse, the register, integration of databases, criterion, method of the analysis of hierarchies, emergency situation, the failure, potentially dangerous object.