

A Research based on Comprehensive Evaluation of Web-Oriented Data Model Design

Ashif Ali

Ph.D. Research Scholar
Department of Computer
Science
C. S. J. M. University,
Kanpur
ashifali76@gmail.com

Dr. Rashi Agarwal

Faculty of Computer Science
Department of Computer
Science
C. S. J. M. University,
Kanpur

Dr. Renu Jain

Head of the Department
Department of Computer
Science, C. S. J. M.
University, Kanpur

***Corresponding Author:** Ashif Ali, Ph.D. Research Scholar, Department of Computer Science, C. S. J. M. University, Kanpur, ashifali76@gmail.com

ABSTRACT

The web has evolved from static information representation to dynamic interaction, making it difficult to find relevant resources without capturing the meaning or semantics. Web data modelling techniques capture and translate complex web applications into easily understood representations, ensuring high data quality and reducing development time, maintenance time, and redundancy. There are three basic levels of data models: Conceptual, Logical, and Physical. A semantic web data model integrates vast amounts of online data and requires customer concern for describing inherent semantics. A requirement framework is needed to capture requirements and generate an equivalent conceptual model in an automated way. The conceptual level design for web data models offer a better way to express problems and increase understand ability from the end user's perspective. However, a strategic approach for implementing the conceptual model into a physical level database schema is needed. Object Management Group's Model-Driven Architecture (MDA) provides an open, vendor-neutral approach to separate business and application logic from the underlying technology platform. Platform Independent Model (PIM) exhibits a specified degree of platform independence of data declaration, while Platform-Specific Model (PSM) is derived from PIM by transformation(s) using necessary rules for describing platform-specific details. Performance evaluation is crucial for data model performance, with correctness, completeness, adaptability, and understands ability being major parameters. Techniques for semantic data model performance evaluation have been proposed, but they are mostly focused on evaluating semantic words or semantic queries. This paper proposes a conceptual level web data model to understand complex relationships within entities, aiming to express existing relations through abstractions and efficiently embed semantics within them. The Semantic Graph Data Model (SGDM) is a multi-layered hierarchical graph with a two-tier approach, representing complex relations among constructs through an object-oriented approach.

Keywords: Computer Architecture, System Security, AI model, Web development.

INTRODUCTION

The web has evolved from static information representation to dynamic interaction, making it difficult to find relevant resources without capturing the meaning or semantics. Semantic representation aims to provide meaningful descriptions of resources that improve the discovery, integration, and reuse of information. Web data modelling techniques capture and translate complex web applications into easily understood representations, ensuring high data quality and reducing development time, maintenance time, and redundancy. There are three basic levels of data models: Conceptual, Logical, and Physical. A semantic web data model integrates vast amounts of online data and requires customer concern for describing inherent semantics [1]. A requirement framework is needed to capture requirements and generate an equivalent conceptual model in an automated way. The framework should be adaptive to cope with changes in requirements and reflect them in the corresponding domain model, reducing the development effort of a web-based information system [2].

The conceptual level design for web data models offer a better way to express problems and increase understand ability from the end user's perspective. However, a strategic approach for implementing the conceptual model into a physical level database schema is needed. Object Management Group's Model-Driven Architecture (MDA) provides an open, vendor-neutral approach to separate business and application logic from the underlying technology platform. Platform Independent Model (PIM) exhibits a specified degree of platform independence of data declaration, while Platform-Specific Model (PSM) is derived from PIM by transformation(s) using necessary rules for describing platform-specific details [3].

A conceptual level representation should be supported by a platform independent and platform-specific representation for a holistic approach to database design. The correctness of a model transformation is crucial for the success of a proposed data model. The dissertation aims to research further into web data modelling, linking ideas from previous attempts to build a conceptual level web data model, especially for better semantic representation. The goal is to automate the requirement gathering process and provide a complete feasible data modelling solution for real-life web-based applications [4].

MAJOR OBJECTIVES AND FINDINGS

This dissertation focuses on identifying issues in the context of semantic web data models and their performance evaluation. The focus is on capturing user concern in the modelling phase, as semantics can be domain-specific and depend on the user's concern. Aspect-orientation is used to capture the user's concern for a specific activity or action. To get the correct result of a semantic search or query, the conceptual web data model must capture the semantic definition [5]. The integrated requirement framework is essential for Model-Driven Development (MDD), as it helps identify and treat functional requirements during the requirement gathering phase. Existing techniques to requirements engineering provide good support for identification and treatment of requirements, but they do not explicitly support domain model development for a specific business application in a semi-automatic fashion. The conceptual model needs to be mapped to the physical level implementation, specifically the PIM and PSM of the existing conceptual model. A strategic approach for implementation should be included within the scope to offer an end-to-end "doable solution." Performance evaluation is crucial for data model performance, with correctness, completeness, adaptability, and understands ability being major parameters. Techniques for semantic data model performance evaluation have been proposed, but they are mostly focused on evaluating semantic words or semantic queries. The information preservation of the proposed web data model is essential during different model transformations towards realization [6].

LITERATURE SURVEY

This paper reviews existing work on conceptual data modelling approaches, physical level implementation efforts, requirement specification automation initiatives, and performance evaluation measures. Classical approaches for data modelling in web are discussed, while advanced web application design demands the concern of customers to be modelled for describing inherent semantics in an integrated way. Aspect orientation provides an approach to capture the respective concern of a user. Physical level implementation is also important as it specifies the necessary mapping for a specific platform as needed by the customer. Existing approaches for conversion of the conceptual model into physical level database schema design are addressed [7].

The evolution of hypertext web data models and existing approaches for developing web applications have been studied. The Hypergraph Data Model was proposed for page-oriented databases in 1989, providing a formal description of database state based on directed labelled hypergraphs. The emergence of XML (eXtensible Markup Language) as the standard

for data representation and exchange on the web brings further necessity of a web data model that can cater XML elements. A new object-oriented data model named XOM was proposed that provides a common model for both hypertext and XML data. Web Modeling Language (WebML) was proposed in 2005, using the Entity Relationship (ER) model for describing data structures. Attila proposed a new modeling approach in 2005 based on Unified Modeling Language (UML) for small and medium scale web application design. The Graph Data Model (GDM) proposes an alternative data modeling to offer higher level abstraction through semantic grouping of constructs. Current web applications consist of highly heterogeneous mix of data such as web pages, XML data elements, and files. A suitable Dataspace Support Platform (DSSP) is proposed to offer several services over dataspace, including search, query, categorization, indexing, and some forms of information recovery [8].

PROPOSED SYSTEM

This paper proposes a conceptual level web data model to understand complex relationships within entities. The model aims to express existing relations through abstractions and efficiently embed semantics within them. The semantic of a word is domain-specific, and user concerns play a key role in determining the actual semantic. The Semantic Graph Data Model (SGDM) is a multi-layered hierarchical graph with a two-tier approach, aiming to express underlying semantics while maintaining object-oriented features [9]. The model is represented by a graph $G(V, E)$, with vertex and edge sets containing various types of nodes and edges. The semantic web-based information system (SGDM) is a model that conceptualizes real-life problems in the e-retail system. T1 represents complex relations among constructs through an object-oriented approach, with two types of vertices: Object Node (ON) and Class Node (CN). Class Node (CN) is a higher level abstraction made over ON, and multiple levels of CNs can be defined through specialization [10].

The primary key of a CN is defined as "determinant" and represented as filled circle arrow-head. Three types of edges are proposed in T1, including Solid Directed Edge with Label, Solid Directed Edge without Label, and Broken Directed Edge without Label. T2 is built on top of T1, expressing the inherent semantic of the problem domain. It includes constructs like Aspect Node (AN), Aspect Link, Page Node (PN), Page Links, Semantic Edge, Ontology Node (OTN), Ontology Link, Directed Ontology Edge, and Ontology Interpretation Node (OIN). The Ontology Management Layer (OML) is introduced in T2 to capture the definition of different semantic terminology involved in a business decision for a

particular domain [11]. The OML consists of four constructs: Ontology Node (OTN), Ontology Interpretation Node (OIN), Ontology Link, Directed Ontology Edge, and Class Node. The label used in the edge expresses the concern on which the term is being interpreted. A Class Node may have multiple aspects based on various user concerns [12].

REQUIREMENTS & ANALYSIS

The first step in Model-Driven Development (MDD) is requirement gathering and analysis to ensure the project's feasibility. This is crucial as requirements are updated frequently by clients, affecting the development project and requiring time, resource, and cost involvement [13].



Fig.-1 Development of FB1 for SGDM T1

The proposed conceptual model, SGDM, should be backed up by a requirement framework that provides a semi-automatic conversion of user requirements to the domain model for a typical business application. This domain model could be considered a Platform Independent Model (PIM) for SGDM conceptual model, as it could be implemented in any database platform.

The SGDM requirement framework proposes a mechanism for automated (up to a certain extent) development of the domain model from the requirement using a novel approach called "e-synthesis". This semi-automatic domain model generation helps accommodate changes within the model according to changes in requirements, reducing the development effort of a web-based information system and ensuring better design of a technical solution [14].

The user requirement gathering consists of two parts: Entity Specification and Relationships. The Entity Specification Template captures requirements for the "e-retail system" and includes attributes like specialization, activities with other entities, and business rules. The Relationships table stores different Parsing Strings, their relevant symbols, and relationship constraints. The automatic generation of the domain model is done based on the facts present in the Fact Base (FB). The Basic User Input Template is used to provide users with semantic queries or statements to execute on a desired domain model based on SGDM. These queries should be simple sentences using semantic keywords based on existing entities provided by the user. The Semantic Keyword Realization Table is used to identify semantic keywords and their relationships [15].

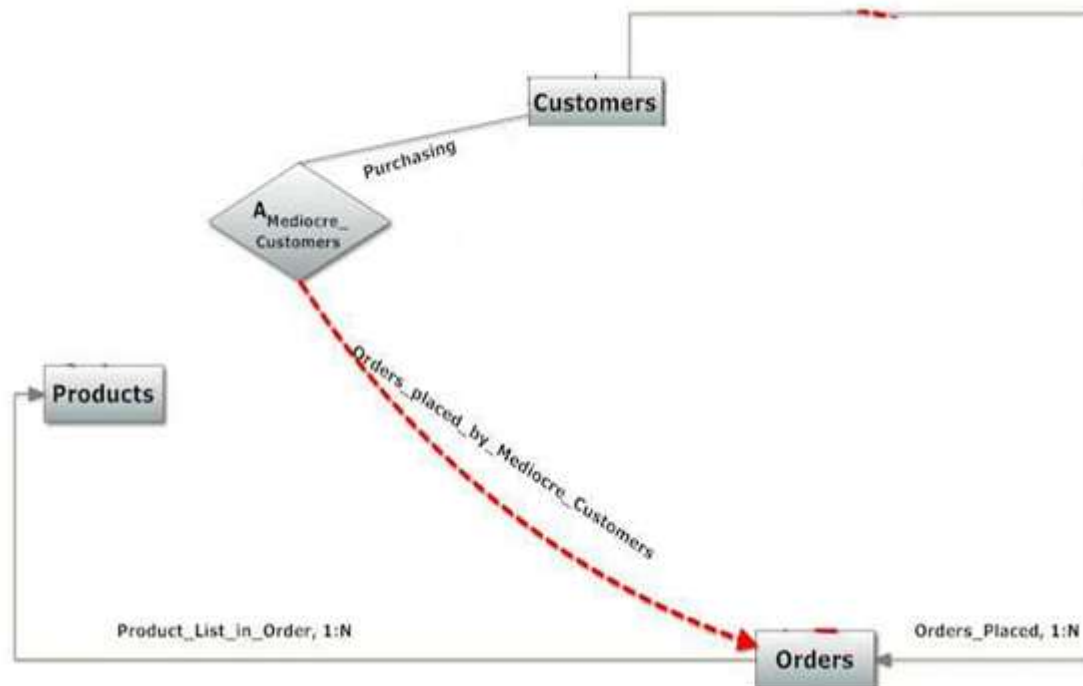


Fig.-2 Resultant sub-graph of SGDM

The Construct Conversion Table stores the mapping rule for domain model design directly from the requirement given in the Basic User Input Template. The table includes columns for entity specialization, Linked Entity Activity Rules, and Concern Keys for identifying correct ontology definitions from OML for a specific AN. This helps in understanding the equivalent constructs of SGDM for domain model design [16].

ARCHITECTURE & IMPLEMENTATION

The text introduces the use of the Generalized Modeling Environment (GME) for creating domain-specific models of large-scale systems. GME is a model-integrated program synthesis tool that allows users to define new modeling languages using UML-based meta-models [17]. It is based on Modeling Integrated Computing (MIC) concepts and is a Configurable Domain-Specific Modeling Environment (CDSME) developed by the Institute for Software Integrated Systems (ISIS), Vanderbilt University. The case study of the "e-retail system" is implemented using GME version 4.11.10. The corresponding interpreter for generating equivalent SQL is done using GME compatible VC++ 6.0. The vocabulary of domain-specific languages implemented by different GME configurations is based on a set of generic concepts built into GME itself. The basic elements of GME include folders, Atoms, Models, Aspects, and a set of constraints. The choice of these generic concepts is the most critical design decision [18]. The GME supports various concepts for building large-scale, complex models. In GME, connections, references, sets, attributes, and constraints are used to express relationships between objects. Connections can be directed or undirected, and they can have attributes themselves. References are similar to pointers in object-oriented programming languages and must appear as part of a model to establish a relationship between the model and the referred-to object. Sets can specify a relationship among a group of objects, with the only restriction being that all members must have the same container and be visible in the same Aspect [19]. Attributes can be text, integer, double, Boolean, and enumerated. Constraints are validity rules applied to a model and can be checked online or in on-demand mode. GME uses a meta-modeling paradigm to configure GME for creating meta-models, which are then automatically translated into GME configuration information through model interpretation. The meta-model is first drawn using the constructs of the SGDM, checked against constraints, and interpreted to generate a new paradigm.xmp file [20]. The domain model is then drawn based on the conceptual model and interpreted with a user-defined interpreter, usually written in Visual C++.

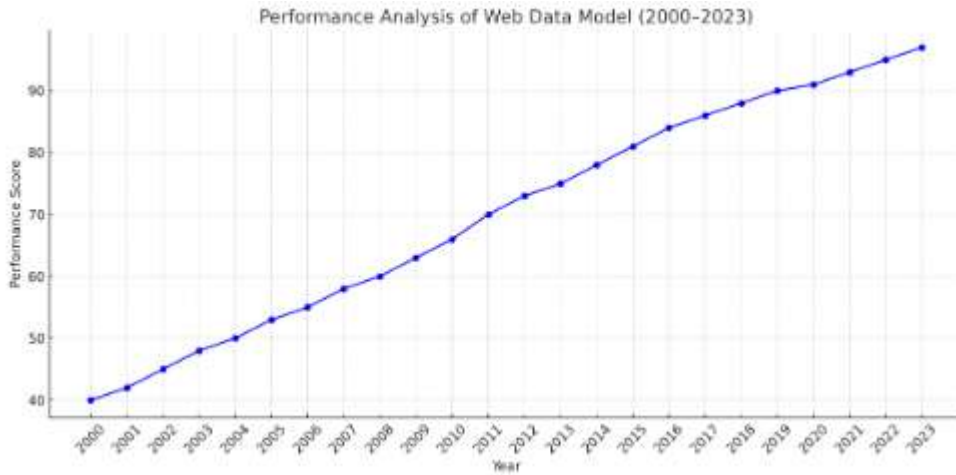

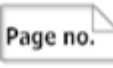






Fig.-3 Analysis graph

SGDM using GME

Table-1 Representation of Mapping Rules of SGDM Constructs with GME Notations

SGDM Constructs	Graphical Notation	GME Meta-Model Notation	GME Meta-Model Notation
CN		Model	
ON		Atom	
Broken Directed Edge without Label		Connection	
Broken Edge without Label		Connection	
Directed Edge with Label		Connection	
Directed Edge without Label		Connection	
Attributes		Atom	

Determinant	Primary Key 	Atom	<code><<Atom>></code>
Page Node	Page no. 	Atom	<code><<Atom>></code>
Page Link		Connection	<code><<Connection>></code>
Ontology Node	Mediocre 	Model	OTN <code><<Model>></code>
Ontology Interpretation Node	Purchasing Capacity (C) \$400< C <<\$1000 	Model	OIN <code><<Model>></code>
Ontology Directed Edge	concern 	Connection	Association_by_Concern <code><<Connection>></code> Degree : enum

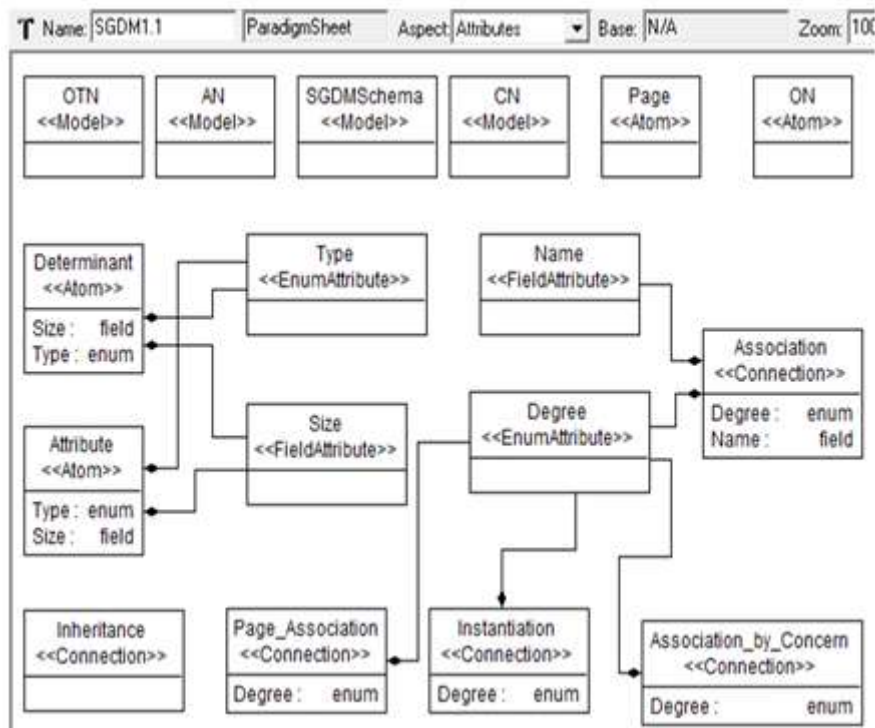


Fig.-4 Constructs of SGDM conceptual model in GME

Table-2 Equivalent OR Schema Generated for SGDM Elements

SGDM Elements	Equivalent OR Schema
CN without inheritance	OBJECT TYPE/TABLE
CN with single inheritance	OBJECT TYPE
CN with multiple inheritances	NESTED OBJECT TABLE
AN	OBJECT TABLE
ON	Tuple of OBJECT TABLE
OTN	OBJECT TABLE
OIN	OBJECT TABLE



Fig.-5 Database creation in Oracle

This paper demonstrates the efficient implementation of a semantic web data modeling (SGDM) data model for an e-retail system using GME tool. The domain model is developed

from user requirements using an integrated framework, and the SGDM Interpreter is developed using VC++ language for database schema generation.

CONCLUSION

This research proposes a graph-based data model for web applications that can model complex relations through hierarchical layered structures. The model can express semantics considering user concerns and interpret domain-specific terms with intended semantics. The model offers a pictorial representation for high understanding and features of the object-oriented paradigm. The logical level implementation of the proposed semantic data model is described using Object Specification Language (OSL), while the physical level implementation is provided by automatic conversion of the conceptual model into a Product Specific Model (PSM). A requirement framework is also proposed for semi-automatic generation of domain models based on user requirements. The methodology for generating an end-to-end solution is described through an evolutionary approach, using Triple Graph Diagrams to ensure model consistency.

References

1. Straub DW, Watson RT. Research Commentary: Transformational Issues in Researching IS and Net-Enabled Organizations. *Info Syst Res* 2001;12(4):337-345. [doi: 10.1287/isre.12.4.337.9706]
2. Fink D, Nyaga C. Evaluating web site quality: the value of a multi paradigm approach. *Benchmarking* 2009;16(2):259-273. [doi: 10.1108/14635770910948259]
3. Markaki OI, Charilas DE, Askounis D. Application of Fuzzy Analytic Hierarchy Process to Evaluate the Quality of E-Government Web Sites. In: *Proceedings of the 2010 Developments in E-systems Engineering*. 2010 Presented at: DeSE'10; September 6-8, 2010; London, UK p. 219-224. [doi: 10.1109/dese.2010.42]
4. Aranyi G, van Schaik P. Testing a model of user-experience with news websites. *J Assoc Soc Inf Sci Technol* 2016;67(7):1555-1575. [doi: 10.1002/asi.23462]
5. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess B, editors. *Analyzing Qualitative Data*. Abingdon-on-Thames: Routledge; 2002:187-208.
6. Akgül Y. Quality evaluation of E-government websites of Turkey. In: *Proceedings of the 2016 11th Iberian Conference on Information Systems and Technologies*. 2016

- Presented at: CISTI'16; June 15-18 2016; Las Palmas, Spain p. 1-7. [doi: 10.1109/cisti.2016.7521567]
7. Arrue M, Vigo M, Abascal J. Quantitative metrics for web accessibility evaluation. 2005 Presented at: Proceedings of the ICWE 2005 Workshop on Web Metrics and Measurement; 2005; Sydney.
 8. European Commission (2016), “*Europa web guide*”, The EU Internet Handbook, available at: <http://ec.europa.eu/ipg/> (accessed 17 June 2018).
 9. Bañón-Gomis A, Tomás-Miquel JV, Expósito-Langa M. Improving user experience: a methodology proposal for web usability measurement. In: *Strategies in E-Business: Positioning and Social Networking in Online Markets*. New York City: Springer US; 2014:123-145.
 10. Wang WT, Wang B, Wei YT. Examining the Impacts of Website Complexities on User Satisfaction Based on the Task-technology Fit Model: An Experimental Research Using an Eyetracking Device. In: *Proceedings of the 18th Pacific Asia Conference on Information Systems*. 2014 Presented at: PACIS'14; June 18-22, 2014; Jeju Island, South Korea.
 11. ISO (n.d), “*Terms & definitions*”, Online Browsing Platform (OBP), available at: <https://www.iso.org/obp/> (accessed 18 February 2021).
 12. Jainari, M.H., Baharum, A., Deris, F.D., Mat Noor, N.A., Ismail, R. and Mat Zain, N.H. (2022), “A standard content for university websites using heuristic evaluation”, in Arai, K. (Ed.), *Intelligent Computing. SAI 2022. Lecture Notes in Networks and Systems*, Springer, Cham, Vol. 506, doi: 10.1007/978-3-031-10461-9_19.
 13. Król, K. and Zdonek, D. (2020), “*Aggregated indices in website quality assessment*”, *Future Internet*, Vol. 12 No. 4, p. 72.
 14. Law, R. (2019), “*Evaluation of hotel websites: progress and future developments*”, *International Journal of Hospitality Management*, Vol. 76, pp. 2-9.
 15. Morales-Vargas, A., Pedraza-Jiménez, R. and Codina, L. (2020), “*Website quality: an analysis of scientific production*”, *Profesional de la Información*, Vol. 29 No. 5, p. e290508.
 16. Thielsch, M.T. and Hirschfeld, G. (2019), “*Facets of website content*”, *Human-Computer Interaction*, Vol. 34 No. 4, pp. 279-327.

17. Whitenon, K. (2021), “*Triangulation: get better research results by using multiple UX methods*”, Nielsen Norman Group, available at: <https://www.nngroup.com/articles/triangulation-better-research-results-using-multiple-ux-methods/> (accessed 4 March 2021).
18. Allison R, Hayes C, Young V, McNulty CAM. Evaluation of an educational health website on infections and antibiotics: a mixed-methods, user-centred approach in England. JMIR Formative Research. 2019 doi: 10.2196/14504. (forthcoming) [DOI] [PMC free article] [PubMed] [Google Scholar]
19. B. Naveen, J. K. Grandhi, K. Lasya, E. M. Reddy, N. Srinivasu and S. Bulla, "Efficient Automation of Web Application Development and Deployment Using Jenkins: A Comprehensive CI/CD Pipeline for Enhanced Productivity and Quality," 2023 *International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*, Erode, India, 2023, pp. 751-756, doi: 10.1109/ICSSAS57918.2023.10331631.
20. "IEEE Recommended Practice for the Internet - Web Site Engineering, Web Site Management and Web Site Life Cycle," in *IEEE Std 2001-2002 (Revision of IEEE Std 2001-1999)*, vol., no., pp.1-114, 3 March 2003, doi: 10.1109/IEEESTD.2003.94235.