ijamrsd.com ISSN: 2583-6404
Journal of Advanced Multidisciplinary Research Studies and Development
Peer Reviewed Referred Journal Bi-monthly Open Access

MATHEMATICAL MODELING OF CONFIDENTIAL INFORMATION SPREAD IN THE DIGITAL ERA

Rama Vijaykumar
Assistant Professor, Department of Mathematics, S.I.W.S. N.R. Swamy College of Commerce & Economics and Smt. Thirumalai College of Science (Autonomous), Wadala, Mumbai, 400031, India.
Email: ramav@siwscollege.edu.in

Abstract
In the modern digital era, confidential information can spread rapidly through social media platforms, messaging applications, and online networks, often leading to important privacy cracks and security threats. Understanding the dynamics of such propagation is essential for predicting spread patterns and designing effective control policies. This paper presents a mathematical modeling framework based on systems of differential equations to analyze the transmission of confidential information among online users. A compartmental structure comprising Susceptible (S), Exposed (E), Infectious (I), and Removed (R) individuals is used to capture main stages of information exposure and sharing behavior. The model incorporates the key parameters such as contact rate, activation delay, and moderation effectiveness. Analytical results, including the basic reproduction number R0, provide insights into the conditions under which confidential information spreads widely or dies out. Numerical simulations are conducted using the Euler method to approximate the solutions of the ODE system and to study how variations in parameters affect the speed and magnitude of information spread. The findings provide insights into technological intervention strategies, such as rate-limiting, awareness enhancement, and rapid moderation, thereby contributing to enhancing information-governance policies. The proposed differential-equation model serves as a useful tool for researchers and digital-platform administrators to analyze and mitigate the risks associated with unintended/ uncontrolled confidential information transmission in the digital era.
Keywords: Confidential Information Spread, Digital Era, Differential Equations, SEIR Model, Mathematical Modeling, Numerical Simulation, Social Media Transmission.

► *Corresponding Author: Rama Vijaykumar*

Introduction:

In today's interconnected digital era, confidential information can spread rapidly across online platforms, social networks, and messaging services. This uncontrolled propagation poses security and privacy challenges. To analyze and predict the dynamics of such spread, Epidemic mathematical models using systems of differential equations are formulated.

In today's highly interconnected digital era, the rapid spreading of confidential information has become a critical concern for individuals, organizations, and governing bodies. Platforms such as social networks, instant-messaging applications, and online communities enable information to travel across vast user populations within a seconds. While this interconnectedness develops communication, it also increases the liability of confidential information to accidental leakage, unauthorized sharing, and deliberate exploitation. Once released, confidential information can propagate uncontrollably, leading to privacy violations, financial losses, reputational damage, and large-scale security threats.

Understanding how such information spreads across digital ecosystems requires systematic and quantitative modeling approaches. Classical studies often draw parallels between information spread and the transmission of infectious diseases, as both involve contact-driven propagation, stages of exposure, and the possibility of controlling or slowing the spread. This analogy inspires the construction of compartmental models where individuals transition between behavioral states based on their interactions and awareness levels.

To capture the dynamics of confidential information spreading, this study adopts a differential equation based SEIR type framework that divides the online population into four key groups: Susceptible (S), Exposed (E), Infectious (I), and Removed (R). Each compartment represents a different phase in the information division cycle, from initial unawareness to active spreading and eventual end due to moderation, awareness, or deliberate non-participation. The mathematical structure allows the analysis of fundamental quantities such as the basic reproduction number R_0, which indicates whether the confidential information is likely to go viral or decline.

Alongside theoretical analysis, numerical simulation plays a vital role in understanding complex behaviors of the system. **Euler's Method** is implemented to approximate the solutions of the differential equations and explore how variations in key parameters - contact rate, activation delay, and moderation effectiveness affect the trajectory of information spread. These simulations reveal threshold behaviors, peak exposure times, and the influence of intervention strategies.

Given the growing importance of digital privacy and data protection, mathematical modeling offers a valuable tool for platform administrators, cybersecurity experts, and policy makers. By quantifying the conditions under which confidential information multiplies, the proposed model contributes to the development of targeted mitigation strategies such as forwarding limits, verification steps, awareness campaigns, and rapid takedown mechanisms.

Review of Literature:

Mathematical modelling is a versatile tool that can be applied across diverse fields; it is not limited to traditional mathematical problems and can be used to study phenomena such as diseases, rumors, artificial intelligence, cyber security and many other real-world systems.

Mamta Kumari (2025–26) analyzed the spread of India's Unified Payments Interface (UPI) by developing a simplified compartmental ODE model aligned with the vision of Viksit Bharat. The model [6] divides the population into potential users, active users, and inactive or dropout users. Sensitivity analysis showed that UPI's long-term sustainability depends on three key parameters: increasing the adoption rate, reducing user dropout, and enhancing reactivation of inactive users. The study highlights how mathematical modelling can support digital-policy design by identifying quantitative levers that strengthen user engagement and ensure a stable, resilient digital payment ecosystem. Studies of the Mandovi River [7] show growing pollution from municipal waste, casinos, and the hospitality sector, with CSIR-NIO and GSPCB reporting increased micro plastics and heavy metals. As natural regulatory processes become strained, differential-equation models help analyze pollutant transport. Similar issues in the Ganga, Yamuna, Lake Tai, and Charles River underscore the need for such modelling, motivating this compartmental study.

Vijaykumar R. (2025) presents a clear analysis of tuberculosis dynamics using compartmental models, including latent infection. The review covers key modelling studies, with accurate formulation, stability analysis, and interpretation of R_0. Simulations show the impact of treatment and vaccination [8]. Greater parameter justification and discussion of limitations would further strengthen the work.

The rapid diffusion of information through digital platforms has motivated extensive research into the mechanisms that govern online spread. Early studies on information propagation drew parallels with classical epidemic models, recognizing that the transmission of ideas, rumors, and digital content resembles the spread of infectious diseases. Daley and Kendall (1964) pioneered the **DK model** for rumor propagation [1], introducing compartments representing ignorant, spreader, and suppressed populations. Their foundational approach established the relevance of compartmental models in understanding social communication processes.

Subsequent research expanded this perspective by adapting epidemiological frameworks such as **SIR** and **SEIR models** to digital environments. Bettencourt et al. (2006) examined social contagion through deterministic differential equations and highlighted the importance of threshold behavior in information diffusion [2]. Similar adaptations were made by Nekovee et al. (2007), who proposed network-based models for the spread of rumors and misinformation, integrating contact rates and individual susceptibility within online communities [11].

With the growth of social media platforms, researchers increasingly focused on understanding virality and the mechanisms driving rapid information sharing. Study by Borge-Holthoefer et al. (2012) demonstrated that digital information often exhibits multi-stage transmission patterns, including awareness, hesitation, activation, and cessation stages—patterns consistent with SEIR dynamics [9]. These works emphasized the role of latent periods, which parallel the exposed (E) class in epidemiological models, thereby validating the use of multi-compartment differential-equation frameworks for digital information spread.

In the context of **confidential or sensitive information**, limited but growing literature highlights the risks associated with uncontrolled dissemination. Research by Chen, Wang, and colleagues (2019) investigated the propagation of sensitive data leaks in online networks and showed that user behavior, platform forwarding limits, and moderation policies significantly influence spreading intensity. Models incorporating user hesitation analogous to the exposed class have proven particularly effective in predicting early-stage growth and eventual saturation [5].

Numerical simulation techniques, including Euler's method and Runge–Kutta schemes, have been widely employed to approximate solutions of nonlinear differential systems used in information diffusion studies. Works by Hethcote (2000) and Li & Liu (2020) emphasized the importance of numerical approximations for exploring parameter sensitivity, outbreak thresholds, and control interventions [4] [11].

Overall, existing literature establishes a strong foundation for applying SEIR-based differential equations to study the dynamics of digital information spread. However, few studies focus specifically on the diffusion of **confidential information**, which involves distinct behavioral, contextual, and ethical considerations. This research contributes by extending SEIR modeling to the digital-privacy domain, integrating hesitation effects, active spreading behavior, moderating influence, and numerical simulations to better understand and mitigate confidential-information propagation.

Basic ODE Model:

We consider a closed population N of online users. At time t, the population of online users be divided into four compartments:

- S(t): Susceptible individuals who are unaware of the confidential information.
- E(t): Exposed individuals who have received but not yet shared the confidential information.

- $I(t)$: Infectious individuals who are actively spreading the confidential information.
- $R(t)$: Removed individuals who stop Infectious individuals spreading (due to awareness, moderation, or policy enforcement).

The total population of online user is given by

$$N = S(t) + E(t) + I(t) + R(t)$$

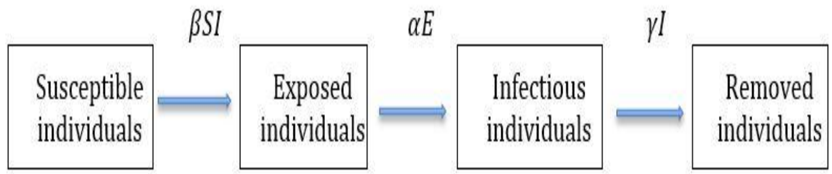


Figure 1: SEIR model flow for Confidential Information Spread in the Digital Era

The model is governed by the following system of ODEs:

$$\frac{dS}{dt} = -\beta SI \qquad (1)$$
$$\frac{dE}{dt} = \beta SI - \alpha E \qquad (2)$$
$$\frac{dI}{dt} = \alpha E - \gamma I \qquad (3)$$
$$\frac{dR}{dt} = \gamma I \qquad (4)$$

where:

- β: contact rate of information transmission.
- α: rate at which exposed individuals become spreaders.
- γ: rate of removal or cessation of spreading.
- N: Total number of online users.

The following assumptions are considered in this study:

- The total population of online users is assumed to remain constant throughout the modelling period.
- Every infected online user is assumed to have an equal probability of being removed.

SEIR Model:

S: number of susceptible users
E: number of exposed users
I: number of infected users
R: number of removed users

$$\frac{dS}{dt} = -\beta SI \qquad (1)$$
$$\frac{dE}{dt} = \beta SI - \alpha E \qquad (2)$$
$$\frac{dI}{dt} = \alpha E - \gamma I \qquad (3)$$
$$\frac{dR}{dt} = \gamma I \qquad (4)$$

Add equation (1), (2), (3) and (4) to get

$$\frac{dS}{dt} + \frac{dE}{dt} + \frac{dI}{dt} + \frac{dR}{dt}$$

$$= -\beta SI + \beta SI - \alpha E + \alpha E - \gamma I + \gamma I = 0$$
$$= \frac{d}{dt}(S + E + I + R) = 0$$
$$S(t) + E(t) + I(t) + R(t) = \text{constant}$$

For initial value t=0, constant =N we get, S(t) + E(t) + I(t) + R(t)= N
(We often take N=1 for normalized proportions.)
Each term has a direct interpretation: βSI is the incidence (new exposures per unit time), αE is the flow from exposed to infectious, and γI is the removal flow.

Model Analysis:
R_0 is the expected number of secondary infections produced by a single infectious individual introduced into a fully susceptible population. For SEIR (where only III transmits), we can compute R0 using the next-generation matrix (infected compartments: (E, I)).
Write the new-infection vector and the transition vector for infected compartments:

$$F=(\beta SI\ 0\ ),\quad V=(\alpha E\ \gamma I - \alpha E\ )$$

Linearize at the disease free equilibrium (DFE) where $S=S_0$ (usually $S_0=1$), E=I=0. Compute the Jacobians with respect to (E, I):

$$F=\frac{\partial F}{\partial (E,I)}|_{DFE} = (0\ \beta S_0\ 0\ 0\ )\ ,\quad V=\frac{\partial V}{\partial (E,I)}|_{DFE} = (\alpha\ 0\ -\alpha\ \gamma\ )$$

The next-generation matrix is $K=FV^{-1}$. Compute V^{-1}

$$V^{-1} = \frac{1}{\alpha\gamma}(\gamma\ 0\ \alpha\ \alpha\ )$$

$$\text{Now } FV-1 = \frac{1}{\alpha\gamma}(0\ \beta S_0\ 0\ 0\ )(\gamma\ 0\ \alpha\ \alpha\ ) = \left(\frac{\beta S_0}{\gamma}\ \frac{\beta S_0}{\gamma}\ 0\ 0\right)$$

The spectral radius (dominant eigenvalue) of K is $\frac{\beta S_0}{\gamma}$. Thus

$$R_0 = \frac{\beta S_0}{\gamma}$$

If we normalize $S_0=1$, then $R_0 = \frac{\beta}{\gamma}$

Interpretation: R_0 depends on the contact/virality β and the average infectious period $\frac{1}{\gamma}$. The exposed class E delays transmission but does not change the number of secondary cases per infective in a fully susceptible population (it only affects timing).

Stability of the Disease-Free Equilibrium (DFE):
Linearizing the full system around DFE and using standard results of next-generation theory yields:
- If $R_0<1$: the DFE is **locally asymptotically stable**. Small introductions of the confidential item die out on average.
- If $R_0>1$: the DFE is **unstable** and an epidemic (widespread information propagation) can occur.

Thus $R_0=1$ is the threshold for take-off.

Peak Condition and Epidemic Dynamics:
Although the closed SEIR system (without demographics) does not support a nontrivial endemic equilibrium with I > 0 as t→∞ (eventually I→0), there is a transient epidemic peak. A useful

instantaneous condition for when the number of spreaders stops increasing (i.e., $\frac{dI}{dt} = 0$) can be derived from $\frac{dI}{dt} = \alpha E - \gamma I$,

and the exposed equation at quasi-steady behavior, one finds that the instantaneous sign of $\frac{dI}{dt}$ changes when the susceptible fraction S reaches

$$\beta S = \gamma \Longrightarrow S = \frac{\gamma}{\beta} = \frac{1}{R_0}.$$

So $S = \frac{\gamma}{\beta}$ is the susceptible level at which the infectious population reaches its peak (analogue of the SIR peak condition).

Final Size (total fraction exposed by the end):

For closed populations the final size relation for SEIR is the same as for SIR with identical transmission and removal rates: the exposed compartment acts as a delay but does not change the final fraction infected. Denote S_∞ the susceptible fraction as $t\rightarrow\infty$. Then $S\infty$ satisfies the transcendental equation

$$ln\frac{S_\infty}{S_0} = -R_0(1 - \frac{S_\infty}{S_0})$$

(when normalized S0=1, $lnS_\infty = -R_0(1 - S_\infty)$). This relation is obtained by integrating $\frac{dS}{dR} = \frac{-\beta SI}{\gamma I}$ and eliminating time; the derivation is identical to the SIR derivation because only I transmits.

Non-dimensionalization (brief):

Scale time by infectious time $\frac{1}{\gamma}$. Introducing dimensionless parameters yields

$\underline{\beta} = \frac{\beta}{\gamma}, \underline{\alpha} = \frac{\alpha}{\gamma},$ and the dimensionless R0= $\underline{\beta}$. This highlights that the important control knobs are the ratios $\frac{\beta}{\gamma}$.

Interpretation & Policy Implications:

- **Reducing** β (platform friction, de-ranking, rate limits) lowers R_0 and may prevent take-off.
- **Increasing** γ (faster moderation, takedowns, legal notices) shortens the infectious period and reduces R_0.
- **Changing** α affects timing: a smaller α (slower activation) delays the epidemic peak and may give moderators more time; it does not change R0 directly.
- **Detect & act early:** because peak occurs when S falls to $\frac{1}{R_0}$, early interventions that keep S near 1 (prevent initial exposures) have the largest impact.

Numerical Simulation & Sensitivity Analysis:

Numerical methods such as Euler's method and Runge–Kutta (RK4) can be used to approximate solutions of the ODE system. Simulation experiments allow us to test the effect of changing all the parameter on the dynamics of spread.

We simulated the SEIR model using Excel with normalized initial conditions:

S(0) = 990, E(0) = 5, I(0) = 5, R(0) = 0

Here, S(t), E(t), I(t), and R(t) denote the fractions of susceptible, exposed, infected, and removed online users, respectively, satisfying the constraint S(t) +E(t) + I(t) +R(t) =1000 for all t.

For the simulation, the following parameter values are used:

$\beta = 1.8, \alpha = 0.5, \gamma = 0.4$

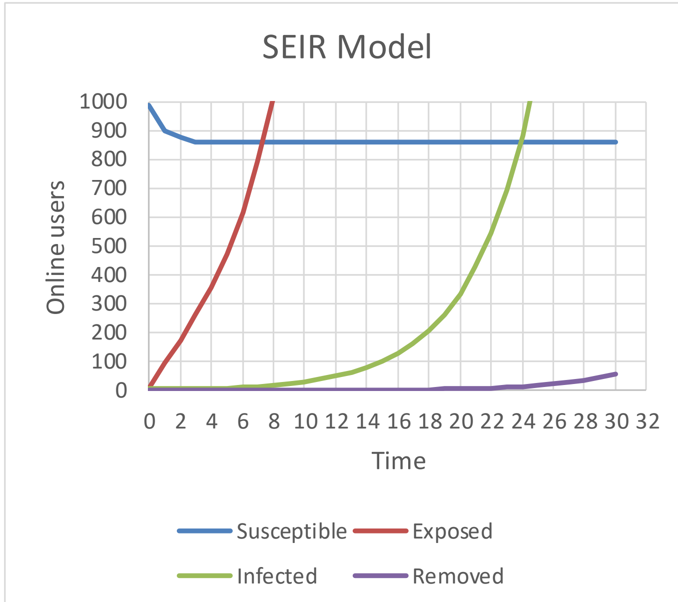


Figure 2: SEIR model simulation of Confidential Information Spread in the Digital Era

Interpretation:

The exposed class E(t) peaks first, followed by the infectious class I(t). This occurs because users must first receive the information and enter the exposed stage E(t) before they can begin spreading it and transition into the infectious stage I(t).

Thus: E peaks first□I peaks later, This time delay is controlled by α.

In the **SEIR** model:

- S → E : A user receives confidential information.
- E → I : After some delay, the user decides to share it.
- I → R : They eventually stop sharing (moderation, awareness, fear, loss of interest).

This leads to the natural timing: People must first be exposed before they can start spreading.

Let us see the behavior of all the four compartments:

I. S(t) = Users who have not yet received or seen the confidential information.

Mathematical equation : $\frac{dS}{dt} = -\beta SI$

S(t) curve starts high (e.g., 99% of users unaware), decreases smoothly as more users become exposed and eventually approaches a **low value** when almost everyone has been exposed. Mathematically the term βSI is always positive, so $\frac{dS}{dt} < 0$ which implies S(t) always decreases over time. So S(t) does **not** have a peak — it is **monotonically decreasing**.

Real world meaning of S(t) curve: S(t) falling means more and more people are becoming aware of the confidential information and the potential audience for *new* spread reduces.

II. E(t) = Users who have *seen the confidential information* but are *not yet forwarding it*.

Mathematical equation: $\frac{dE}{dt} = \beta SI - \alpha E$

They hesitate, read but don't react, verify information, wait before forwarding.

At the beginning, many susceptible users (S(t)) exist. Spreaders (I(t)) expose them quickly therefore, the term βSI is large → E(t) rises sharply.

E(t) peaks when: $\frac{dE}{dt} = 0 \rightarrow \beta SI = \alpha E$, meaning: The rate of new exposures equals the rate of activation into spreaders.
S(t) decreases sharply as the population becomes aware. Fewer new exposures occur and Individuals in E(t) transition to I(t) so E(t) declines toward zero.
Real-world meaning of E(t) curve: It represents how many users **see but hesitate**. It peaks early because awareness spreads faster than actual sharing thus Indicates the **latent potential** for future spreading.

III. I(t)= Users who are actively sharing/forwarding the confidential information.

Mathematical equation : $\frac{dI}{dt} = \alpha E - \gamma I$
Slow Initial Growth as Few individuals start as spreaders, they depend on E(t) to supply new spreaders, Since E(t) is still rising, I(t) grows slowly.
Rapid Growth After E(t) Peaks: When E(t) is large, the influx α E is high. More exposed individuals activate into spreaders and I(t) increases rapidly.
I(t) peaks when $\frac{dI}{dt} = 0 \rightarrow \alpha E = \gamma I$ which means the rate of becoming a spreader equals the rate of stopping spreading.
As E(t) declines, fewer new spreaders enter. Moderation, loss of interest, or saturation increases R(t) and I(t) eventually falls to zero.
Real-world meaning of I(t) curve: Represents **active virality** of confidential information. I(t) curve peaks later than E(t) because users must first be exposed before they can spread, which indicates the **maximum risk** period for information leakage.

IV. R(t)= Removed/ stopped spreading- Online users who delete the message, stop re-sharing, lose interest or get restricted/moderated.

Mathematical equation : $\frac{dR}{dt} = \gamma I$
R(t) always increases as there is **no exit** from R and γI is always ≥ 0 So $\frac{dR}{dt} > 0$ therefore R(t) rises continuously. R does **not peak**; it is **monotonically increasing**.
Graphical study says that the curve R(t) starts at zero, rises slowly at first, rises rapidly near the peak of I(t) (because γI is largest), rises slowly again once I(t) declines and eventually levels off to a final value.
Real-world meaning of R(t) curve: R rising means, moderation is happening, online users voluntarily stop sharing, digital platforms reduce virality and forwarding fatigue kicks in.

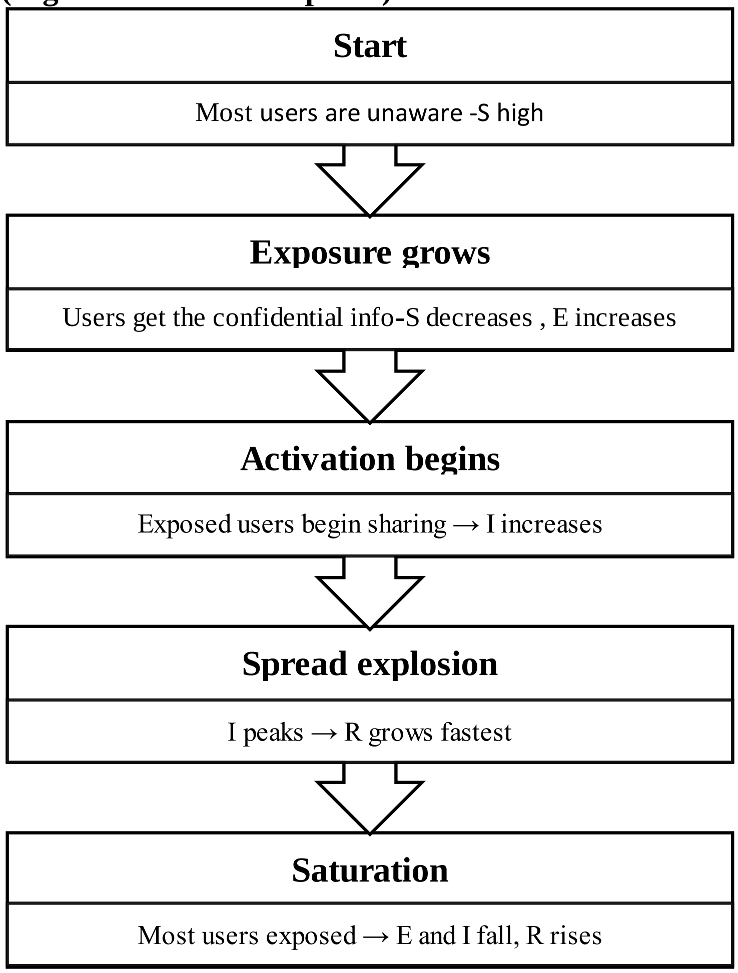
Tabular Representation of Numerical Simulation & Sensitivity Analysis:

Let us see the clear and tabular explanation of how S(t) and R(t) behave in the SEIR model of confidential-information spread in the digital era—along with how they relate to the peaks of E(t) and I(t).

Table 1: Summary for all the four curves S(t), E(t), I(t) & R(t)

Variable	**Meaning**	**Behavior**	**Peak**
S(t)	Unaware users	Always decreasing	No peak
E(t)	Exposed but not sharing	Rises, peaks, then falls.	Peaks first
I(t)	Actively spreading	Rises after E(t), peaks later	Peaks after E(T)
R(t)	No longer spreading	Always increasing	No peak

Intuitive Timeline (Digital Information Spread) Flow Chart:



Key take away is, S decreases because more users become aware, E rises then falls, I rise later then falls, and R increases as people stop spreading, together forming the typical SEIR information-diffusion curve.

Control Strategies:

Potential interventions include:

- Reducing β by limiting forwarding or adding verification steps.
- Reducing α by increasing awareness before individuals share.
- Increasing γ through rapid moderation and takedown policies.

Optimal control theory can be applied to minimize the harmful spread while balancing cost of interventions.

Conclusion:

Differential equation models provide a powerful tool to study the spread of confidential information in the digital era. Such models help in understanding key parameters influencing virality and in designing effective countermeasures.

This study presented a differential equation based mathematical model to analyze how confidential information spreads in the digital era. Using the SEIR framework, the population was categorized into Susceptible, Exposed, Infectious, and Removed classes, capturing the stages of awareness, hesitation, active dissemination, and eventual cessation. The model successfully represents the real-world dynamics of online information flow, where users first receive information, then decide whether to share it, and eventually stop spreading due to loss of interest, moderation, or external interventions.

The analysis revealed that the Exposed class peaks before the Infectious class, highlighting the inherent delay between receiving confidential information and actively sharing it. Numerical simulation using Euler's method further demonstrated how the transmission rate, activation rate, and removal rate significantly influence the magnitude and duration of the spread. Increased virality intensifies the outbreak, while improved moderation and awareness strategies effectively reduce the peak and total spread.

Overall, the SEIR model offers a strong framework for understanding, predicting, and controlling the spread of confidential information online. The insights obtained can assist digital platforms, policymakers, and organizations in designing targeted interventions that limit harmful dissemination while maintaining a safe and secure digital environment.

Bibliography:

1. Daley, D. J., & Kendall, D. G. (1964). *Epidemics and rumours*. Nature, 204, 1118.
2. Bettencourt, L. M. A., Cintrón-Arias, A., Kaiser, D. I., & Castillo-Chávez, C. (2006). *The power of a good idea: Quantitative modeling of the spread of ideas*. Physica A, 364, 513–536.
3. Zhao, L., Xie, Y., Zhang, J., & Wang, J. (2013). *Rumor spreading model considering hesitating mechanism in* social networks. Physica A, 392, 987–994.
4. Hethcote, H. W. (2000). *The mathematics of infectious diseases*. SIAM Review, 42(4), 599–653.
5. Chen, J., Wang, X., & Li, Y. (2019). *Modeling and analysis of sensitive-information diffusion in online social networks*. Physica A, 523, 1209–1220.
6. Kumari, M. (2025–26). A Compartmental ODE Model for Digital Payment Adoption: A Case Study of UPI in India's Journey to Viksit Bharat. GENIUS, 14(1), 61–65. (ISSN: 2279-0489)
7. Kumari, M. (2025) Mathematical Modeling of Casino Waste Impact on the Mandovi River, Proceedings of the International Online Conference on Advancing Innovations Through Mathematics: A Journey Across Disciplines (AITM-2025), Publisher: Department of Mathematics, Bapatla Engineering College ISBN: 978-93-344-0268-1 Pages: 123-133
8. Vijaykumar, R. (2025) Mathematical Modelling for Tuberculosis Dynamics and Control, Proceedings of the International Online Conference on Advancing Innovations Through Mathematics: A Journey Across Disciplines (AITM-2025), Publisher: Department of Mathematics, Bapatla Engineering College ISBN: 978-93-344-0268-1 Pages: 174-183
9. Borge-Holthoefer, J., Baños, R. A., González-Bailón, S., & Moreno, Y. (2013). *Cascading behavior in complex socio-technical networks*. Journal of Complex Networks, 1(1), 3–24.
10. Li, X., & Liu, F. (2020). *Numerical approaches for nonlinear differential-equation models in environmental systems*. Environmental Modelling & Software, 132, 104806.
11. Nekovee, M., Moreno, Y., Bianconi, G., & Marsili, M. (2007). *Theory of rumour spreading in complex social networks*. Physica A, 374, 457–470.