ZURICH

How Do We Navigate a Safe AI Transformation

An insurer's perspective on managing risk in the age of artificial intelligence

Section 1

The transformation we are in

Artificial Intelligence (AI) is reshaping the global economy at a pace that outstrips regulation, organizational adaptation, and public understanding.

Global AI adoption has surged from 20% of organizations in 2017 to near-universal levels in 2025[1], with global corporate investment in AI expected to have reached $258.7 billion in 2025[2]. Nearly 80% of business executives report AI is reshaping work-force strategies[3]; 54% expect significant job-role transformation within two years[4].

For insurers, this is not an abstract trend. We sit at the intersection of AI as both users and enablers, deploying AI within our own operations while underwriting AI-related risks for customers. That dual perspective creates a powerful feedback loop: insights from thinking on how to underwrite emerging AI risks strengthen internal governance, while operational experience improves our ability to assess and price those risks.

We are participants, not spectators

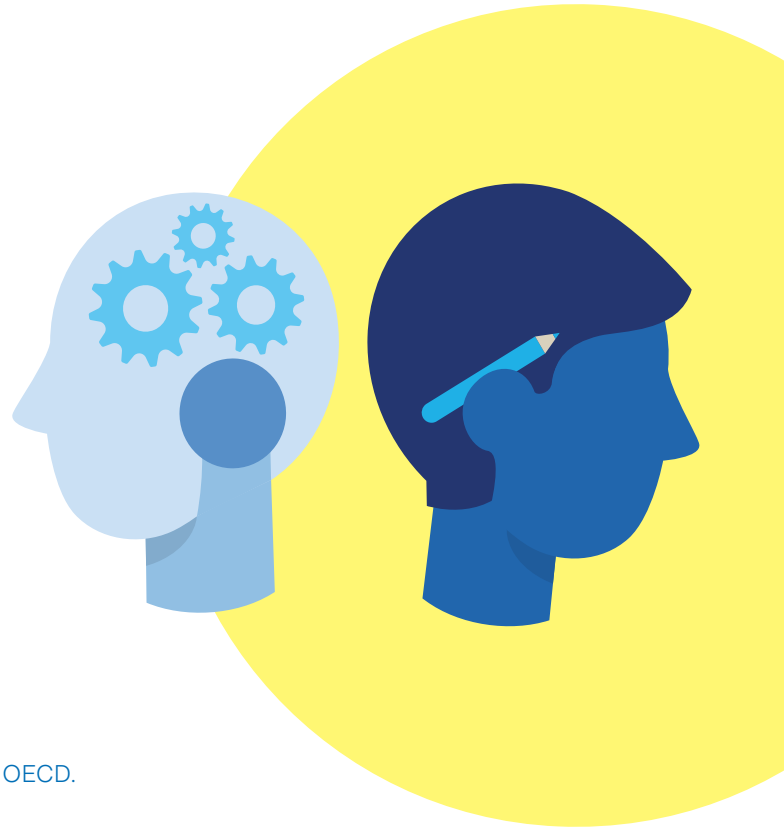
As Zurich advances its digital transformation strategy, the company's technology investments have exceeded $2 billion while operating over 460 Generative AI use cases live in production across the full insurance process.[5] More than 90% of our retail distribution is now digitized. Zurich eXchange 2.0, our suite of APIs handling around 80 million transactions per month, uses Generative AI to fast-track embedded insurance, generate documentation, and enable coverage checks, policy comparisons, and claims insights at scale – all of which help enhance the experience of our customers.

AI in underwriting. AI-driven solutions help under-writers scale their capacity by capturing and reusing underwriting expertise in a group-wide, connected knowledge base that is accessible through natural language.

AI in claims processing. AI agents also handle initial claim reporting across channels, enabling claims reporting at any time.

Commercial insurance and reinsurance are leading adoption, precisely because they combine complex decision-making with scarce specialist capacity: the conditions where AI-driven knowledge democratization delivers measurable impact.

This is part of a broader trend. Across the industry, insurers using AI for claims processing report 40–50% faster handling times and 25% higher customer satisfaction. AI-powered services route routine inquiries automatically and escalate complex issues to human experts with full context, reducing waiting times by 60%[6] .



[1] The State of AI: Global Survey 2025 | McKinsey.
[2] Venture capital investments in artificial intelligence through 2025 | OECD.
[3] The State of AI: Global Survey 2025 | McKinsey.
[4] The Future of Jobs Report 2025 | World Economic Forum.
[5] Annual Report 2025 | Zurich Insurance Group.
[6] Triple-I Blog | Global Insurers Embracing AI for Claims Resolution, Customer Service and 40 Claims Redemption Rate Statistics: Key Facts Every Organization Should Know in 2025.

Yet the opportunity is only half the picture. Alongside developing new AI-enabled products and services, we are thinking carefully about governance, accountability, and data management.

Zurich's technological innovation and engineering experts have developed Agentic AI Principles that set clear guidance for ethical, transparent, and purpose-driven AI deployment. These principles invite every AI solution to be explainable, continuously monitored, and managed with well-defined ownership, ensuring both trust and accountability[7].

Our AI risk framework categorizes AI systems by risk level and encourages enhanced controls for high-impact applications. This includes documented risk assessments, bias and impact testing, strong data stewardship, and structured incident reporting, expecting that regulatory and ethical standards are built into every step of development and deployment.

To reinforce oversight, Zurich applies a 'human-in-the-loop' principle for consequential decisions, i.e., decisions with material outcomes for the individual. Whenever AI influences outcomes such as coverage, claims, or customer interactions, a qualified human would review and validate the final decision, preserving judgement, empathy, and accountability.

We intend to underpin these measures with dedicated structures: the Zurich AI Council strives for strategic oversight, setting policy and priorities; the AI Architecture Board ensures technical and operational rigour. This integrated approach means governance is is built in by design, not bolted on as an afterthought.

Insurance exists to make progress possible by managing risk intelligently. The same principle must guide AI: not risk elimination (forgoing benefits), not risk ignorance (inviting catastrophe), but smarter risk management.

[7] AI at Zurich | Zurich Insurance Group.

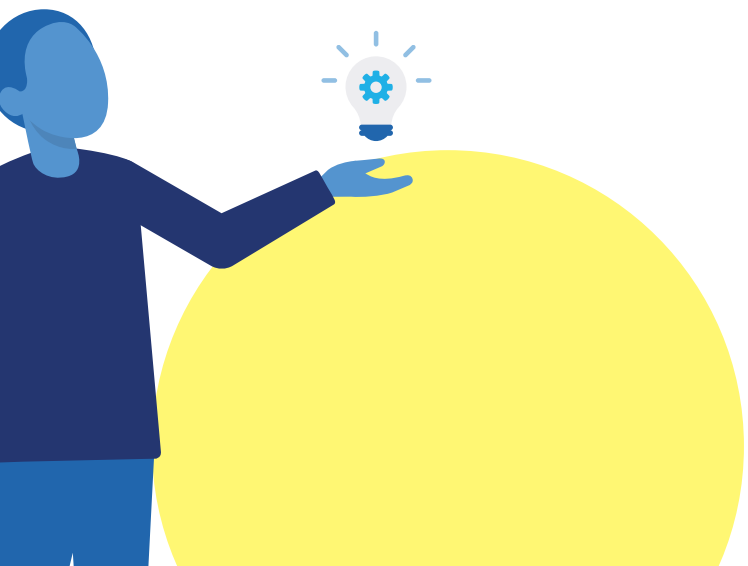
Section 2

The risk landscape

Insurers are among the world's most experienced assessors of risk. Through decades of underwriting catastrophe risk, cyber exposure, and critical infrastructure, we have developed frameworks for thinking about interconnected, fast-moving threats that resist traditional modeling. We believe that expertise can play an important role in helping customers and policy makers better understand risks associated with widespread AI adoption.

Our experience in cyber risk offers perhaps the most direct insight. In Closing the Cyber Risk Protection Gap[8], a joint white paper with Marsh McLennan, we mapped the cyber threat landscape as a risk spectrum: from well-understood, quantifiable exposures through to a frontier where visibility is limited and credible quantification does not yet exist. The frontier is not static; it shifts as data, experience, and analytical capabilities advance. There are risks the private insurance market can absorb, risks that require active mitigation and public partnership to address, and risks that remain beyond commercial insurability.

The approach managing AI risk can be understood using the same lens. Some AI failures are quantifiable, insurable, and manageable through established controls. Others – particularly those involving correlated failures, concentration risk, and cascading systemic effects – sit at the frontier, demanding new approaches and, potentially, public-private coordination.



The AI risk spectrum

We see the AI risk spectrum evolve along three segments:

i. Known and quantifiable risks

AI is amplifying and reshaping familiar risks, but these are challenges we already have the expertise to identify, test for, and mitigate:

- **Bias and exclusion.** AI systems trained on historical data can replicate past discrimination in underwriting or claims. A claims tool may systematically disadvantage certain groups.
- **Unsafe or incorrect outputs.** AI may approve inadequate coverage, overlook critical risk factors, or guide customers toward poor decisions.
- **Data security and privacy.** AI systems are data-hungry, creating pressure to collect broadly and retain indefinitely, conflicting with data minimization principles. AI can infer sensitive characteristics from innocuous data, re-identify anonymized datasets, and blur traditional privacy boundaries.
- **Copyright and Intellectual Property infringement.** AI outputs that replicate identifiable copyrighted content without the rights holder's permission.

[8] The cyber security challenge – and how to address it | Zurich Insurance.

ii. Emerging and harder-to-quantify risks

Drawing on our experience in cyber risk underwriting, there is a parallel dynamic in AI, getting you to a point where the private re/insurance market may not be able to absorb alone the aggregate amount of financial loss:

- **Opacity and lack of explainability.** Many powerful AI systems function as "black boxes." When AI denies coverage or rejects a claim, can the decision be explained? Customers may accept human decisions they disagree with if they understand the reasoning; they will not accept algorithmic decisions particularly if no one understands those.

iii. At the frontier: systemic exposure

This is where our experience as cyber insurer provides the sharpest insight. The most undercussed AI risks are correlation and concentration, and they mirror the dynamics that make catastrophic cyber risk so challenging to model and absorb.

- **Correlated failures.** When organizations deploy identical AI systems, failures tend to cluster under shared conditions. If insurers rely on similar catastrophe models that underestimate a given risk, losses can cascade across the industry. Homogeneity creates fragility.

The AI risk spectrum evolves along three segments

i. Known and quantifiable risks ii. Emerging and harder-to-quantify risks iii. At the frontier: systemic exposure

- **Accountability gaps and supply-chain risk**. In complex AI supply chains involving developers, deployers, data providers, and platforms, attributing liability when harm occurs is not obvious. AI development is modular: organizations use third-party foundation models, external training data, and cloud infrastructure. Compromise at any point propagates throughout the chain, mirroring the supply-chain dependencies we already assess in cyber underwriting.

- **AI-empowered cyber threats.** AI broadens the attack surface. It enables more sophisticated phishing, faster vulnerability discovery, and adaptive attacks that learn from defenses. While AI also empowers defenders, the asymmetry is stark: attackers need one successful breach; defenders must protect all vectors simultaneously.

- **Concentration risk.** AI capabilities are currently concentrated among a handful of foundation model providers. A single compromised model could affect thousands of downstream applications. A single provider's service disruption cascades across industries. In cyber, we already see how single points of failure – as demonstrated by the CrowdStrike incident[9] – can have far-reaching implications. We need to understand how the same risks might play out with AI.

- **Cascading incidents.** Deeply interconnected AI systems create feedback loops. One system's output becomes another's input; failures propagate through dependency chains faster than humans can intervene. The 2010 "Flash Crash", where automated trading systems interacting at high speed produced a 600-point market drop within minutes, illustrates the dynamics[10]. AI systems operating across more domains and with greater autonomy – think of agentic AI – amplify this risk significantly.

[9] 2024 CrowdStrike-related IT outages - Wikipedia

[10] Findings Regarding the Market Events of May 6, 2010: Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues.

What this means for insurability

There are existing liability frameworks such as product liability, professional negligence or contractual allocation which remain broadly fit for purpose, and much AI-related harm may be able to be resolved within them. However, AI's rapid evolution and deployment – particularly with the eruption of agentic AI – is generating scenarios for which those frameworks were not designed . Although organizations, as legal persons, will continue to be responsible for those decisions, multi-agent chains, where autonomous systems could interact and delegate across organizational and jurisdictional boundaries, may create accountability gaps that existing doctrine struggles to close. This could be particularly the case where agents are built on open-source components with varying provenance, limited documentation, and no single point of oversight. In such environments, tracing how a decision was reached, which evidence must be disclosed, and assigning responsibility when it causes harm, becomes a materiallydifferent challenge.

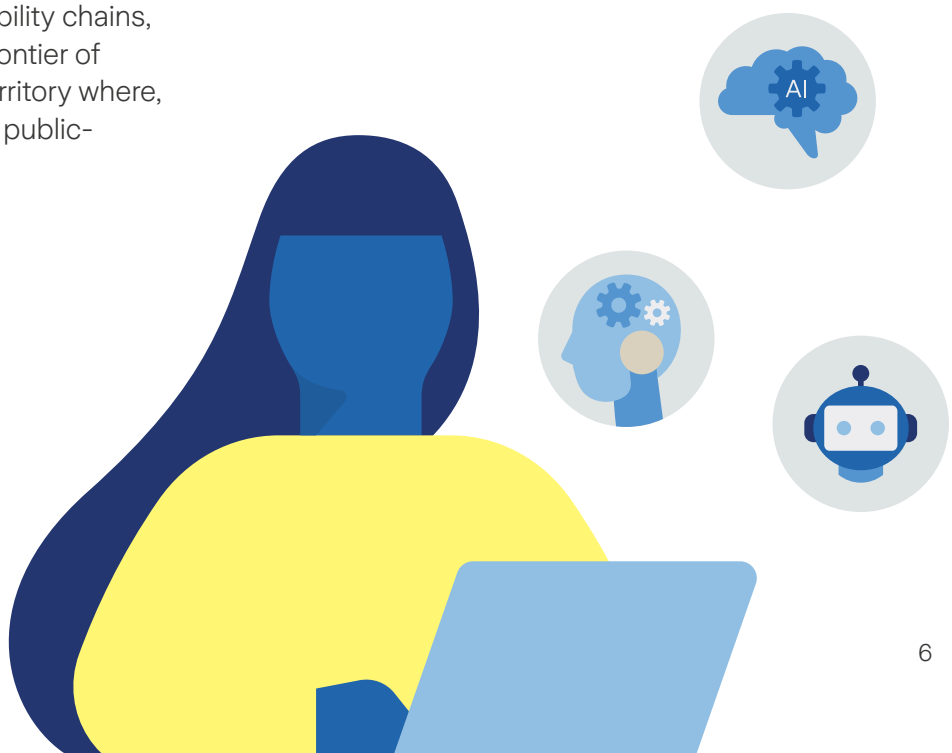
This matters for insurability. We cannot insure what we cannot understand and quantify, or price risk when responsibility is not attributable. Clarity in liability attribution is essential for insurability, and insurance is needed to underpin widespread responsible AI adoption. Without incident data, without mapped concentrations, without clear accountability chains, AI will create risks that sit beyond the frontier of commercial insurability, precisely the territory where, in cyber, we have argued for structured public-private partnership.

From risk analysis to policy response

Policy makers are already converging on a similar reading of the landscape. The prevailing direction is pragmatic, proportionate governance, which leverages existing regulatory frameworks where possible and favouring technology-neutral and outcome-focused approaches. This is the right approach and best exemplified by the UK government's pro-innovation strategy[11]. But the arrival of agentic AI complicates it materially: when autonomous systems interact and delegate with minimal human involvement, the 'human-in-the-loop' paradigm that underpins most current governance thinking will, for many operational steps, simply not hold.

There is also growing recognition that AI, in all its variations, constitutes a form of critical infrastructure, carrying systemic risks that echo those already seen with cyber. The social and economic consequences of a failure of a critical infrastructure requires a focus on resilience and not just recovery. This means not just technical robustness, but the capacity of economies and societies to absorb, adapt to, and recover quickly from AI-related disruption.

Insurers should be well placed to inform policy thinking around this agenda. Modelling complex, interconnected, and fast-evolving risk is what we do. Translating that expertise into practical frameworks for policy makers is a natural extension of the role we already play in cyber, catastrophe, and infrastructure risk. The priorities that follow reflect that perspective.



[11] The UK's Framework for Responsible AI Regulation

Section 3

Policy priorities

The preceding analysis points to three interconnected priorities that will require policy action to address and that will demand coordinated action from industry, government, and civil society.

1. Trust through transparency, accountability, and clear liability

Public trust is the foundation of AI adoption. Without it, capable systems face resistance and regulatory backlash. Mature liability frameworks already cover substantial ground, but AI is evolving faster than their perimeter, and where the boundary of liability becomes unclear, so does insurability.

Establish clear rights to notice when AI is used in consequential decisions, to explanation, to contestation, and to meaningful redress. These must be practical; complex appeal processes are not meaningful protection.

Require transparency about AI use in consequential decisions. A simple, standardized notification could establish a universal baseline: informing people whenever they are interacting with, or subject to decisions influenced by, AI. Where decisions carry material impact, transparency alone is not sufficient. Human review must be preserved and audit trails and execution tracing embedded, as Zurich does through its Agentic AI Principles, so that every AI action is traceable and contestable.

Clarify liability attribution. Insurers cannot price risk when responsibility is unclear. Clear liability enables insurability; insurability enables adoption.

Both regulatory ambiguity and regulatory overreach carry costs. Unclear application of liability in AI contexts forces organizations into defensive over-compliance or leaves genuine accountability gaps unaddressed. Equally, layering AI-specific liability regimes on top of mature existing frameworks risks fragmentation and unnecessary complexity. The priority should be ensuring that established liability frameworks are fit for purpose in AI contexts – through proportionate clarification, for example, for evidence disclosure, and specific criteria – rather than constructing parallel regimes. This would support consistency in judicial interpretation. A functioning AI risk transfer market depends on that balance.

2. Proportionate governance and resilience for systemic risk

Effective governance must scale with risk. At the systemic level, no comprehensive AI incident reporting framework exists, and concentration risk among a few large AI providers remains largely unaddressed.

Adopt risk-based governance that matches requirements and oversight intensity to actual risk. High-risk applications (coverage decisions, claims denial, critical infrastructure) require rigorous bias testing, human oversight, documentation, and third-party audits. Low-risk applications need lighter touch.

Treat significant AI incidents like cyber incidents: mandatory reporting, anonymized information sharing, and collective learning. No comprehensive AI incident reporting framework currently exists.

Address concentration and correlation risk explicitly. Map dependencies on foundation models and major platforms. Consider diversity requirements for critical infrastructure. Develop circuit breakers that limit cascade propagation, just as financial markets have done for algorithmic trading.

Harmonize core standards internationally. Regulatory fragmentation increases compliance costs. Common taxonomies, documentation standards, and testing expectations – modeled, e.g., on approaches like the International Sustainability Standards Board – would create baseline consistency while allowing local variation. Risk that cannot be measured or attributed cannot be transferred. Without incident data and concentration

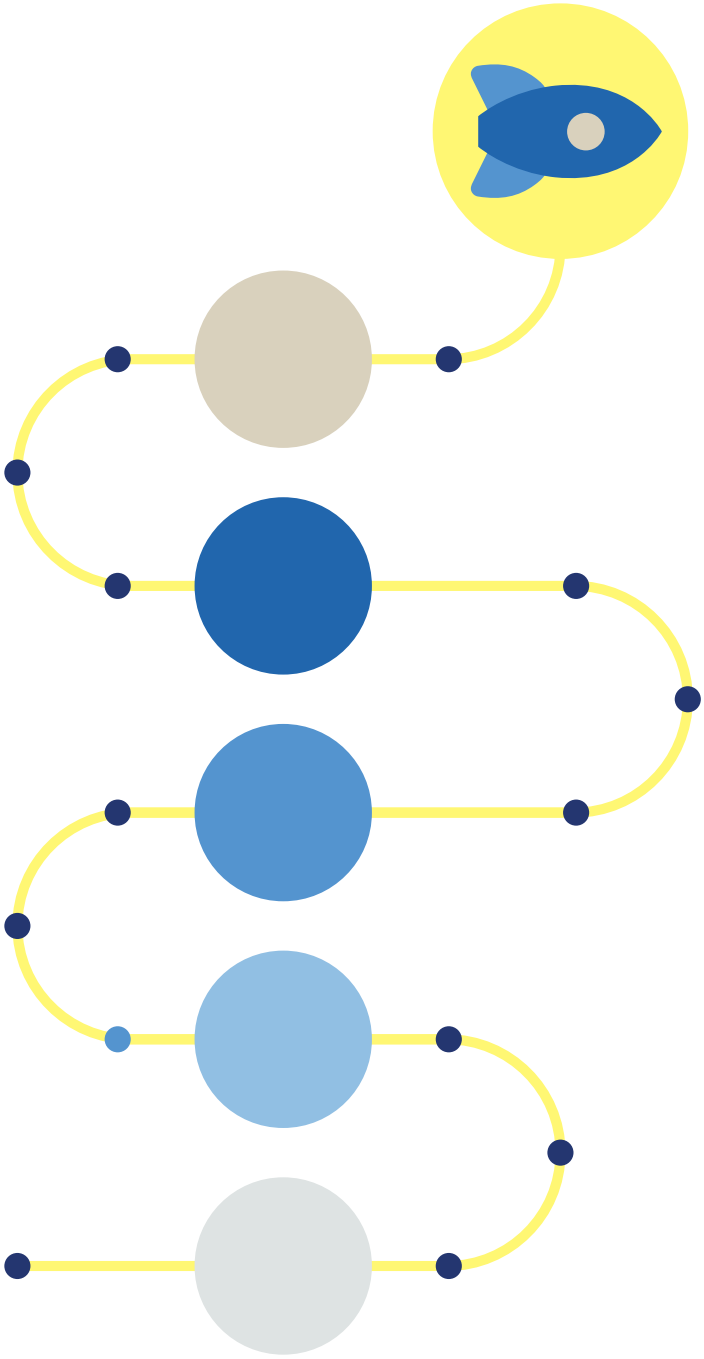
mapping, AI creates silently insured systemic risks – exposure that already sits on balance sheets, unpriced, unmodelled, and unmanaged, the same challenge we have faced at the frontier of catastrophic cyber.

3. Workforce transition and educational adaptation

Insurers are experienced at identifying protection gaps - areas where the risk outpaces the available safety net. We as an insurer of employer liability and workforce-related risks have a grounded perspective also on this challenge. Workforce displacement from AI or poorly managed workforce transitions can generate measurable consequences[12]. AI deployment at pace and scale will amplify these dynamics unless transition support keeps pace.

Invest in workforce transition at scale. Policymakers need to support workforce transition actively. Public-private collaboration is essential. Governments will be held accountable for ensuring tools, training, and opportunities are accessible across demographics and geographies. Employers who commit to workforce development and support building AI literacy will have a competitive edge. Programs like Singapore's SkillsFuture and Germany's or Switzerland's "Kurzarbeit" system offer proven models.

Preserve essential human skills. Beyond reskilling for new capabilities, we all will need to examine which core skills such as ethical reasoning, contextual judgment, relationship building or creative problem-solving remain indispensable and must be retained in education and professional development.



A shared responsibility

AI transformation is underway. The choices we make now about governance, industry practices, and societal investment will determine whether benefits will materialize or stall.

Insurance has always had a key role in supporting innovation and progress by managing risk intelligently. We bring long-standing expertise in assessing systemic threats, pricing uncertainty, and building frameworks that enable responsible risk-taking.

We are committed to leading by example through accountable deployment, lifecycle governance, and transparency while working with policymakers and civil society to build guardrails that are strong enough to prevent harm and flexible enough to accommodate innovation.

The goal is not zero risk. It is risk that is understood, governed, and proportionate to impact — channeled toward a future where AI can continue to be deployed at scale with confidence and trust.

[12] See Zurich's report on 'The value of mental health' examining what mental health conditions mean for people, productivity and protection systems across six markets: Australia, Chile, Germany, Malaysia, the UAE and the UK.

ZURICH

Zurich Insurance Group (Zurich) is a leading global multi-line insurer founded more than 150 years ago, which has grown into a business serving more than 82 million customers in more than 200 countries and territories, while delivering industry-leading total shareholder returns.

Reflecting its purpose to 'create a brighter future together,' Zurich offers protection services that go beyond traditional insurance, to support its customers in building resilience. Since 2020, the Zurich Forest project has supported reforestation and biodiversity restoration in Brazil's Atlantic Forest.

The Group has more than 65,000 employees and is headquartered in Zurich, Switzerland. Zurich Insurance Group Ltd (ZURN) is listed on the SIX Swiss Exchange and has a level I American Depositary Receipt (ZURVY) program, which is traded over-the-counter on OTCQX. Further information is available at **www.zurich.com.**